

Games people play with authors' names

Authorship of a scientific paper is a privilege that is all too easily abused. Attempts to solve the problem with general rules encounter insurmountable obstacles, but individual accountability is unavoidable.

In one respect at least, artists are to be envied by scientists. Artists may lack the benefit of science's capacity for objective certainty and the immediate recognition of worth that usually comes with it; they may lack the technological relevance that stimulates large investment; and they tend to lack the relative security of an employment contract. But they can generally take for granted what for a scientist is an increasingly rare privilege: being the sole originator of his or her creation. A corollary is that the artist is released from one burden in particular: the bitterness that can arise in scientific collaboration. And nothing is more likely to be divisive than the list of a scientific paper's authors. Do we therefore need more rules of conduct?

Few would dispute that researchers have to take responsibility for papers that have their names on them. A senior laboratory figure who puts his or her name on a paper without direct supervision or involvement is unquestionably abusing the system of credit. There have been occasions where distinguished scientists have put their names irresponsibly on a paper that has turned out to contain serious errors or fraud. Rightly, some of them have paid a heavy price. But how deeply does a principal investigator or supervisor have to be aware of details of a paper bearing their names?

Definition

One response might be a general rule that any co-author should be able to take responsibility for the entire content of a paper. Although that seems reasonable at first sight, it will often fly in the face of practicality, as when an identification of a new hominid depends on the bringing together of independent and specialized skills in dating, stratigraphy, morphological comparison and possibly molecular biology. The rule fails where, as is often the case, the skills and experience involved are complementary but not even contiguous, let alone overlapping. That gives rise to another twist to the debate: if the dating, although essential to the conclusions, is done blind — in other words, without reference to the context of the samples — do the daters deserve more than an acknowledgement? Many would say not.

An alternative and less demanding definition, then, of the threshold for authorship might be that anybody with a name on the paper must have contributed something essential and scientifically original in arriving at its conclusions. But what of agreements (or attempts at them) between researchers that an inventor of an innovative new technique should be included in the authorship of a paper that involves the technique but not its originator? It is not unknown in astronomy, for example, for the inventor of an innovative detector to insist that any use of that instrument requires his name to appear on the paper. This issue has inevitably arisen more frequently in an era in which research is increasingly multidisciplinary. For example, agreements have been struck in physics laboratories to the effect that chemists who synthesize photonically interesting materials should appear only on the first physics paper

to emerge and only be acknowledged thereafter. Such agreements are inevitably *ad hoc*.

Authorship causes trouble enough, without the added strife that arises from attempts to attribute a contribution more precisely. There was a vigorous debate recently in the *Nature* editorial office about whether to permit footnotes to papers indicating that authors X, Y and possibly Z made equal contributions to the result. The debate was split along disciplinary lines — most biologists for, physical scientists broadly against. On the principle that *Nature* should try to provide a service rather than a straitjacket, we decided to respond to a genuine demand apparently restricted to the biological community, hoping that physical scientists would not notice and then start worrying about the issue. So far our hope has been more or less fulfilled. But tolerance runs out when footnotes are sought not only by the first two authors but also, absurdly, by the last two as well.

Responsibility

One might speculate that physical scientists differ on this issue either because they are more mature or, more probably, because they are in less cut-throat areas of research. Certainly they seem more readily to opt for a rational if somewhat impersonal approach: alphabetical listing. High-energy physicist collaborations, notoriously extended, may start the listing with one or two key individuals and will then list alphabetically. One worried author in a less hyper-collaborative branch of the physical sciences went so far as to survey three hundred papers in physics, mathematics and chemistry having two authors, to discover that three-quarters of them had names in alphabetical order. His null hypothesis was that one half had made a conscious decision to list themselves alphabetically (and that the practice appears to be universal in mathematics).

Alphabetical ordering reduces scope for debates about credit, but not recognition: one cannot escape the fact that K. Aa, the scientist whose name happens to appear first in the 1996 *Science Citation Index* listing, has a particular advantage in this respect. That, presumably, is why some long-standing collaborations list themselves alphabetically but in rotation.

There are places where quantification of research assessment has run amok, to the extent that individuals are even rated according to their position on author lists. Such schemes are fatally undermined not least by the fact that particularly distinguished authors may put themselves last on some occasions, penultimately when they wish to give a colleague a leg-up by donating the supposedly privileged last slot, or even first if, selfishly, they can enforce their will over more deserving colleagues when a paper is expected to be much cited.

The conclusion? Attempts to produce even quite simple rules about the deployment of authors' names are doomed to fail at the inevitable emergence of counter-examples. Nevertheless it remains the case that having one's name on a paper must be an acceptance of personal responsibility, a betrayal of which can damage people's trust not only in an individual but also in science itself. □



Congress and Varmus in clash over 'illegal' embryo research

[WASHINGTON] Harold Varmus, the director of the US National Institutes of Health (NIH), came under heavy fire from a congressional subcommittee last week because of lapses that had allowed an NIH-funded researcher to conduct human embryo research in alleged contravention of US law.

"It appears that some in NIH believe they are above the law. They are wrong," said Joe Barton (Republican, Texas), chairman of the oversight and investigations subcommittee of the House of Representatives Commerce Committee, which oversees management of NIH. Barton accused NIH of "lax oversight, miscommunication and an utter lack of sensitivity to the social and moral issues involved".

But Varmus, the only witness at the hearing, said that "We too have a very large respect for law... This episode was highly unusual and unlikely to recur".

The hearing focused on the case of Mark Hughes, a reproductive biologist with whom NIH severed its relationship last autumn on finding that he had breached a congressional ban on federal funding for human embryo research (see *Nature* 385, 190; 1997).

Hughes is a specialist in single-cell genetic analysis, which can be applied both to non-embryonic cells, such as cancer cells, and to early-stage embryos. DNA from 8-cell embryos produced at *in vitro* fertilization (IVF) clinics is analysed and, if it is found to be abnormal, the embryos are not implanted. This technique is called preimplantation genetic diagnosis (PGD).

Hughes, who did not testify at the hearing, said in a subsequent statement that he had felt that the federal ban did not to apply to his work, because he was working with DNA

from embryos, not the embryos themselves.

"I studied and tested only DNA samples obtained from embryos by technicians at non-NIH funded programs at clinics and universities around the country," he said in a written statement. "It was my belief that this in no way violated any prohibitions against government-sponsored embryo research." The statement adds: "Many people at NIH... knew of my single cell work. I think this demonstrates that I believed, in good faith, that I was permitted to continue my PGD research."

Varmus maintained at the hearing that NIH recruited Hughes in 1994 only to do non-embryonic work, unless an NIH policy then in place forbidding embryo research was changed. He said NIH believed that Hughes was doing embryo research only in his own time and with his own resources, at a separate facility, while doing non-embryonic work at the National Center for Human Genome Research (NCHGR, now the National Human Genome Research Institute (NHGRI)), to which he was recruited by NIH from Baylor University in Houston, Texas. Varmus said that Hughes was repeatedly told of existing prohibitions on embryo research. "He knew what the rules were," Varmus said.

But the politicians produced NIH documents that appeared to contradict Varmus. For instance, the agreement Hughes signed with NCHGR in 1994 states, "pre-implantation genetic diagnosis will also be provided by blastomere biopsy and single-cell [polymerase chain reaction] technology". And an NCHGR publicity book outlining the work of its lead researchers describes Hughes' work as including "genetic diagnosis of the 8-cell preimplantation embryo *in vitro*".



Hot seat: Varmus (left), seen here at an earlier hearing on cloning, says events surrounding the controversial research are "unlikely to recur".

Varmus told the politicians that the book was "without any obvious implication that that research was being done at the NIH". He added that, when Hughes was hired, NIH had a "reasonable expectation" that the existing prohibition on human embryo research would soon be lifted. The documents simply reflected that expectation. But Ron Klink (Democrat, Pennsylvania), the senior subcommittee Democrat, declared, "We are very troubled by this. It appears to be considering something" that was forbidden. He demanded that Varmus produce earlier communications to Hughes proving that "there was not a wink and a nod" when Hughes was hired.

The congressmen also distributed a recent report by an internal NIH inquiry committee on protection of human subjects. It concluded that preimplantation diagnosis "was a prominent feature of Dr Hughes' activities, as well as those of at least two [NIH-supported postdoctoral fellows] and one Visiting Fellow throughout their association with NHGRI".

The report says that at least one case of preimplantation diagnosis was handled on the NIH campus. It adds that "the work of the Hughes laboratory was presented at in-house NHGRI research retreat days and in the published abstracts of the American Society of Human Genetics in 1995 and 1996. Thus, it appears to the committee that the research activities of the Hughes laboratory were conducted and communicated in an open manner within the NHGRI and general genetics communities."

The federal law pertaining to Hughes' particular embryo work was clear from January last year, when Congress banned federal funding for research on spare embryos from IVF clinics in which embryos are harmed or destroyed.

But when Hughes was hired in 1994, a legal obstacle to federally funded human embryo

US attacks EU gene-food labelling move

[PARIS] The United States last week attacked European Commission proposals to label genetically modified foods saying they are contrary to free trade (see *Nature* 384, 502; 1997). Raising the prospect of a trade war, Dan Glickman, the US secretary for agriculture, said that America would "not tolerate" segregation of modified bulk products — such as maize — from their traditional counterparts, which he described as impracticable.

This hard line reflects that of the World Trade Organization (WTO), which has ruled that countries cannot refuse imports or require labelling unless they can either show that the product does not meet international standards or produce scientific evidence of

risk (see *Nature* 384, 301; 1996). "Sound science must trump passion," declared Glickman, who argued that genetically modified foods are safe.

The US position could, if not modified, make a trade war inevitable. The 'mad cow' crisis has made European consumers suspicious of scientific reassurances about safety. As a result, Europe is likely to face up to any US challenge at the WTO. Even if the United States won its case, it would probably face a backlash from consumers and retailers against US goods.

The commission adopted the labelling proposal as a directive last week, and it should become law in the European Union by the end of July. □

research had recently been lifted, and NIH was operating, according to Varmus's testimony, with the expectation of revising its own policy to reflect this and allow some embryo research. An NIH expert panel — which included Hughes — concluded in December 1994 that federally funded human embryo research was allowable in certain carefully controlled conditions.

While President Clinton immediately blocked an NIH panel recommendation that some creation of human embryos for research purposes be allowed, his order did not address work on 'spare' IVF embryos. This was banned thirteen months later by Congress.

A document provided by NIH but not produced by the congressmen at the hearing supports NIH's version of events. Jeffrey Trent, the scientific director at NCHGR, wrote to a colleague in August 1995 that he had talked to Hughes about his non-NIH embryonic work at nearby Suburban Hospital. "I told [Hughes] that I am *not* willing to place federal personnel, etc., at Suburban within an IVF facility," Trent wrote.

The NIH also provided letters from Hughes assuring NIH that he was not using NIH-funded staff or equipment at Suburban Hospital — statements later shown to be false, according to an NIH investigation.

At the hearing, Klink also asked Varmus about "all these postdocs that were doing the same thing (that is, preimplantation genetic diagnosis) under Hughes". Did they not "give a damn", he demanded, or were they "intimidated" by Hughes? "I am concerned about a deficiency in mentorship" by Hughes, Varmus responded.

The charge that Hughes intimidated his postdocs into doing work they suspected might be illegal is contained in notes made by Kate Berg, an NIH official, when NIH investigated Hughes in October last year. But Scott Gant, a lawyer representing Hughes, said that Hughes "denies having intimidated anybody."

The NIH human subjects inquiry committee also concluded that Hughes conducted the embryo work at NIH without the required approval of an NIH ethics board, and without meeting the standards for commercial diagnostic laboratories required by law.

Varmus said that NIH was taking corrective actions to reduce the risk of a recurrence, including ensuring that no other similar research was being conducted, and contacting researchers most likely to stray into the area banned by law.

Other management gaffes surfaced during the hearing. The congressmen grilled Varmus about how Hughes was allowed to load a lorry with more than \$1-million-worth of NIH equipment, including some boxes marked "Suburban Hospital" — on loan to Georgetown University — and leave NIH with it, without a completed loan agreement or inventory. Some of the equipment was diverted to Suburban Hospital. **Meredith Wadman**

NIH may drop special funds for 'new investigators'

[WASHINGTON] Officials at the US National Institutes of Health (NIH) are expected to decide next week whether to eliminate a category of grants for new investigators worth on average \$70,000 a year for five years — about half the standard NIH grants.

The abolition of the grants, known as R29s or First Independent Research Support and Transition (FIRST) awards, has been proposed by an NIH working group on new investigators. The working group was co-chaired by Marvin Cassman, director of the National Institute of General Medical Science, and Elvera Ehrenfeld, director of NIH's Division of Research Grants. It was set up last October in response to concern about the impact of tightening NIH funds on young scientists forced to compete against seasoned researchers.

Its report, which is due to be discussed by the directors of NIH institutes on 3 July, recommends that R29 applicants should in future be required to apply for standard NIH grants (R01s), for which competition is fiercer. But their applications would explain that they were from young scientists who were first-time applicants.

Peer reviewers would be told to accept less preliminary information in such applications than they would from established investigators. And grants to new investigators would generally be made for five years; standard grants usually last three years.

Anne Thomas, a spokeswoman for Harold Varmus, the director of NIH, said that approval of the proposal is not a forgone conclusion and that the discussion at the meeting is expected to be substantive.

The working group argues that R29s deliver too little money to make up for the relative advantage to young scientists of applying in a less competitive pool. (In 1995, R29 applicants had a 28.5 per cent chance of winning an award, as against an 18.2 per cent chance for first-time R01 applicants.)

According to the draft report, the relatively small size of R29 awards has grown "increasingly onerous" for recipients, most of whom are forced to seek supplementary funding. "It seems curious that we accept so easily much smaller awards to new investigators who are likely to be less well equipped to deal with such constraints," says the report.

It also points out that those awarded R29 grants are consistently less successful in subsequent grant applications than new investigators whose first grant is a standard one.

The draft report concedes that, while the suggested changes would ensure more money for successful young scientists, "there is clearly a concern that success rates will



drop". It says that it is "essential", therefore, for the NIH to define a desired entry rate for new investigators, and to ensure that this is at least sufficiently high to replace those retiring, about 9 per cent per year.

Ehrenfeld points out that the data analysed by the working group, for the years 1980 to 1995, show that younger scientists are not being selectively disadvantaged by dwindling research dollars. Although grant application success rates have gone down, "they've gone down for everybody", she says. New investigators "are faring as well as anybody else."

The report shows that the success rate for new R01 applicants dropped from 26.2 per cent in 1980 to 18.2 per cent in 1995, while the rate for experienced investigators dropped from 30.5 to 22.5 per cent.

Some scientists are concerned about the possible impact of the changes. John Moore, of the Aaron Diamond AIDS Research Center in New York, says that the impact of R29 awards in his own laboratory has been helpful. He says that the working group has addressed "the critical issue" that a young investigator cannot generate the same amount of preliminary data as an established scientist.

But, reacting to the proposal that new applicants should be judged differently for R01 applications based on identifying information, Moore, a veteran of NIH study sections, says he would ask "would study sections pay much attention to that?"

Others, such as Jais Lingappa, a cell biologist and R29 grant applicant at the University of California, San Francisco, are worried that the changes could mean that fewer young scientists will win funding. But she praises the idea of granting newcomers larger awards, provided they are made for five years. "What the young investigator really needs is three or four years of protected time," she says. **M.W.**

Accord could lift block on European biotech patents

[MUNICH] The prospects for a clarification of Europe's rules on the patenting of transgenic animals and plants brightened last week when the European Parliament's legal affairs committee approved — with amendments — a revised directive drawn up by the European Commission in Brussels.

The way is now paved for approval of the directive — which covers biotechnological inventions and allows the patenting of human genes, as well as of transgenic animals and plants — in the parliament's plenary session next month. An earlier draft directive was rejected by the parliament two years ago (see *Nature* 374, 103; 1995).

Approval is expected despite a vociferous minority of members of the parliament, particularly those representing 'green' parties, who oppose the patenting of life on moral grounds. Others argue that the draft directive is important because it clarifies the European Patent Convention, the book of rules for the European Patent Office (EPO) which was drawn up before the advent of genetic engineering.

The convention explicitly excludes from patentability animal and plant 'varieties' — in order to protect plant breeders' rights — although animals and plants are patentable. This ambiguity has led to a two-year halt by the EPO in the issuing of patents relating to animals and plants that are the products of biotechnological inventions.

In 1995 an EPO board of appeals ruled against the issuing of a patent covering a procedure producing herbicide-resistant

plants and the plants and seeds arising from the process to the Belgian company Plant Genetics Systems. The board argued that a transgenic plant represents a collection of plant varieties (see *Nature* 381, 178; 1996).

In consequence, the EPO has had to put on hold more than 1,100 applications for patents covering plants, and nearly 500 patents covering animals, until the law is clarified. It has put on hold any ruling in the appeal against its most high-profile animal patent application, the Harvard oncomouse.

The new wording in the draft directive, as amended by the legal affairs committee, states that "inventions which concern plants or animals may be patented if the practicability of the invention is not technically confined to a particular plant or animal variety".

This wording clearly distinguishes species and varieties, says Christian Guggerell, an expert on biotechnological inventions at the EPO. If used as a guide by the EPO, even though it is independent of the commission, it would allow the backlog of applications to be processed, he believes.

But first the EPO requires new case law of its own. The next case to be considered by an EPO board of appeals will probably concern an application brought by the company Ciba-Geigy now part of Novartis.

The board of appeals, which could rule as early as next autumn, could refer to the more precise wording on plants and plant varieties in the commission's draft directive to help it to interpret its own more imprecisely written rules.

Alison Abbott

Japan reaches a compromise on organ transplants

[TOKYO] Long-awaited legislation designed to provide a legal framework for organ transplants has at last been passed by Japan's parliament, the Diet. But only after a bizarre compromise was reached on the legal definition of death.

For more than a decade, Japan's medical policy-makers have been trying to gain public acceptance of "brain death" as death of an individual, so that organs such as the liver and heart can be removed from brain-dead patients for use in transplants (see *Nature* 318, 591; 1985). Japan's first heart transplant in 1968 led to demands that the doctor responsible be tried for murder. No heart transplants have been carried out since.

The legislation that passed the Diet last week is a modification of an earlier bill passed by the Diet's lower house in April which allows organs for transplantation to be removed from "brain death bodies" (see *Nature* 387, 5; 1997). But the original bill met strong opposition in the upper house, where many legislators were unwilling to recognize brain death as death.

In the compromise legislation, brain-dead patients are recognized as being dead only if they have given prior written permission for donation of their organs. Otherwise, their heart must stop beating before they are considered to be dead.

Furthermore, as in the original bill, the family of the deceased must give permission for a transplant to be carried out, and the family's wishes will override those of the brain-dead patient.

There are various explanations as to why the Japanese are so opposed to accepting brain death and organ transplants. These include a strong belief that the body and soul are inseparable and a Confucian concept that the body is a gift of one's parents that cannot be given away. Nevertheless, this has not stopped many Japanese from seeking transplants overseas, even in underdeveloped countries such as the Philippines.

The legislation, while opening the door to transplants, is unlikely to lead to widespread organ transplantation, commentators say. Quite apart from an unwillingness to recognize brain death, Japanese people generally are unwilling to donate their organs. Evidence for this can be seen in the fact that kidney transplants, which can be carried out after the heart has stopped beating, are uncommon.

There is also no well-established system of donor cards in Japan. Furthermore, as donors must give prior written consent to the use of their organs, none will become available for children.

David Swinbanks

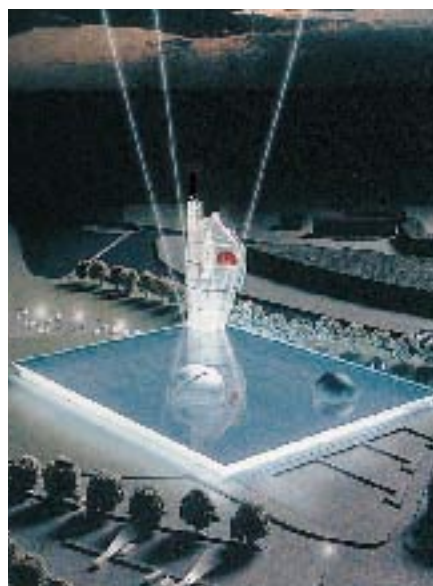
Leicester centre is set for take-off

[LONDON] A UK National Space Science Centre (right) should lift off in Leicester in 2000 with the aim of attracting 300,000 visitors a year, thanks to a grant of £23 million (US\$38 million) from the Millennium Commission.

The centre will include displays and hands-on activities to convey current knowledge about space, astronomy and the global environment. There will be a planetarium and a Challenger Learning Centre — an educational initiative backed by the US space agency NASA — offering simulated space flights.

Exhibits will include objects from the European Space Agency and pieces of moon rock from NASA.

The centre will be built on the site of a decaying sewage treatment works which is owned by Leicester City Council and Severn Trent Water.



Rio follow-up faces frustrated ambitions

[NEW YORK] When delegations from 200 countries met this week to survey progress in the five years since the Earth Summit in Rio de Janeiro, they faced a situation of growing global disarray, with inaction on key issues and divisions between and among rich and poor nations.

The meeting, at the United Nations in New York, was dominated by public pressure on the United States to join Europe in setting dates and percentage targets for reducing emissions of greenhouse gases in advance of the meeting on this issue in Kyoto, Japan, in December.

Tony Blair, the new UK Prime Minister, said: "We in Europe have put our cards on the table. It is time for the special pleading to stop and for others to follow suit." Blair had just returned from a summit meeting in Denver at which US President Bill Clinton again refused to commit the United States to emissions targets.

At the 'Earth Summit+5' meeting, the UN's members were seeking to agree wording for a political statement and a longer "programme of outcomes" to meet the goals agreed at Rio. But the opening of the meeting was marked by a widespread recognition that implementation of these goals has so far been a spectacular failure.

Greenhouse gas emissions continue to grow, in defiance of a voluntary agreement between the developed countries to return them to 1990 levels by the year 2000. Deforestation, desertification and the depletion of fish stocks continue unabated.

The Global Environment Facility — the only major institution whose formation was



Blair: 'the special pleading must stop'.

pledged at Rio — remains far too small to make an impact on development in poor countries, and development aid from the rich countries has fallen sharply.

Ismail Razali, Malaysia's ambassador to the UN, who presided over the summit, opened it by describing the actions taken since Rio as "paltry". He said there was a "recession of spirit" afflicting the signatories to Agenda 21, the main document produced at Rio. "Damning statistics show that we are heading further away from, and not towards, sustainable development."

Non-governmental organizations echoed this acknowledgement of failure. "Five years on, we are extremely disappointed with what has happened on both the environmental and development sides," said Martin Khor, a Malaysian official of the Third World Network.

"The consensus that we had reached at the Earth Summit has fallen apart," says Barbara Bramble of the US National Wildlife Federation. "There is good reason for the South [developing countries] to be suspicious of all those pious statements you'll hear this week about the need to take action."

As criticism of US intransigence on greenhouse gas emissions mounted, there was little indication of irritation from officials in the Clinton administration, some of whom hope that international pressure will help to persuade the US public and the Congress that action is needed to reduce emissions.

The administration remains divided about how far to go at the Kyoto meeting — and constrained by powerful congressional opposition to cuts in emissions. Both houses of Congress would have to pass laws to implement cuts, and the Senate would have to endorse any international agreement.

Last week, for example, the international trade subcommittee of the Foreign Relations Committee in the Senate, chaired by Chuck Hagel (Republican, Nebraska), held hearings that made clear its opposition to mandatory emissions limits for the United States unless developing countries also face mandatory limits. The subcommittee called Senator Robert Byrd (Democrat, West Virginia) and Congressman John Dingell (Democrat, Michigan) as witnesses, to demonstrate that leading figures in both parties share this opposition.

Dingell and all the senators present — with the notable exception of John Kerry (Democrat, Massachusetts) — attacked mandatory limits. They referred repeatedly to the need for action by China which, they said, would surpass the United States as the world's largest emitter early next century.

More than 60 senators have signed a resolution, drafted by Byrd, that would rule out mandatory limits if they would damage the US economy — or if developing countries do not have them too.

European leaders, even as they press the Clinton administration to propose targets for emission reductions, acknowledge its predicament. "I don't think our problem is with the administration," said Robin Cook, the British foreign secretary. "It is the American public that cannot continue as it is" in its consumption of energy, he said.

But Wim Kok, Prime Minister of the Netherlands, which holds the presidency of the European Union, pointed out that many Americans support action, just as "not every European applauds fuel tax increases".

Kok said he understood that it could make sense for Clinton not to publish targets too long before the Kyoto meeting, to shorten the barrage of attack they will inevitably suffer in the United States.

Colin Macilwain

India may set up genetics advisory panel

[NEW DELHI] Prompted by concern about the potential misuse of genetics research, Indian scientists are seeking the creation of a national bioethics panel to advise on the ethics of research and on the management and use of genetic information.

The proposal for such a panel, to monitor all aspects of research using human DNA and genetic testing, was made at a meeting of geneticists, social scientists, lawyers, economists and philosophers from India and abroad held recently in Goa.

The conference, the first of its kind in India, was organized by the Indian Academy of Sciences and was sponsored by the United Nations Educational, Scientific and Cultural Organization (Unesco), the United Nations University in Tokyo and the Third World Academy of Sciences in Trieste, Italy.

"While genetic studies may provide powerful tools for disease diagnosis and treatment, they are prone to abuses which could threaten society," says Prakash

Tandon, president of the Indian academy. "It is therefore imperative that, as we consider developing genetic research in India, we also work in parallel on the ethical, legal and social implications."

Indian scientists have decided that, while following developments in gene therapy closely, they will hold back at present from experimenting with humans. India would also observe the international embargo on germ line therapy.

The government's Department of Biotechnology has agreed to do the groundwork for setting up the ethics panel. Consisting of individuals of "impeccable integrity and proven competence", the panel would lay down ethical guidelines for genetics research, advise policy-makers and safeguard the rights of human research subjects, says Manju Sharma, the department's secretary. The proposed panel would also develop Indian responses to issues such as patenting.

K. S. Jayaraman

Geologist set to challenge 'creationism' verdict

[SYDNEY] The conflict about the competing claims of evolution and the fundamentalist interpretation of the biblical account of creation is due for another burst of public attention in Australia. Last week, Ian Plimer, the geologist who lost his main legal challenge to the creationist Allen Roberts just over three weeks ago (see *Nature* 387, 540; 1997), lodged an appeal to the Full Bench of the Federal Court.

Plimer's lawyers will claim that Justice Ronald Sackville erred in deciding that Roberts had not acted in trade or commerce in raising funds for "archaeological research" on the supposed remains of Noah's Ark in Turkey.

A campaign to help Plimer with his crippling legal costs has also been launched at a public meeting at which the Anglican Arch-

bishop of Adelaide spoke in his support. Plimer, a former Australian Humanist of the Year, has been campaigning against the teaching of so-called "creation science" in schools.

This week, the world's oldest association of geologists, the UK Geological Society, announced it had made Plimer an Honorary Fellow "for his courageous stand against 'creation science'". Richard Selley of Imperial College, London, writes: "Most geologists ignore creationists, and do not believe that it is worth joining them in serious debate. This is an extremely dangerous attitude and we are already seeing the price that must be paid by adopting it both in the USA and Australia. Professor Plimer is a man of enormous courage, who has put his money where his mouth is."

Peter Pockley

Competition is the spur for German research centres

[MUNICH] The Helmholtz Society (HGF), an umbrella organization that oversees Germany's 16 national research centres, is to set up a system of 'strategy funds' worth DM150 million (US\$87 million) for which the centres will have to apply competitively.

HGF centres are financed jointly by the federal government (90 per cent) and the host state (*Land*, 10 per cent), and have a total annual budget of about DM3 billion. In future, each centre will have to contribute 5 per cent of its institutional budget, which covers running costs, investment costs and salaries, to the new funds.

This money will be redistributed through competition. Centres can apply for research projects worth upwards of DM10 million, limited to three years. In the first round, the centres must submit applications by next January to a senate committee of six scientific and six industrial members. Proposals will then be reviewed by a scientific panel that will include foreign members. Germany's main research grant-giving body, the Deutsche Forschungsgemeinschaft, is to support the committee in its choice of referees. The HGF senate, which consists of members from the political, scientific and industrial arenas, will make its final decisions next spring, and the chosen projects will start in July 1998.

Areas likely to be covered by the new funds include bioinformatics, environmental technologies and superconductors. In addition, from next year they will be used to finance HGF's Young Scientists Programme. The society will advertise jobs for around 500 young scientists in the next three years.

After a meeting last week with Joachim Treusch, the HGF president, Germany's research minister, Jürgen Rüttgers, promised support for the new plans.

Scientists at HGF have broadly welcomed the competition plans, but they are concerned that large interdisciplinary HGF centres with a strong base in applied research, such as the Forschungszentrum Jülich, are likely to benefit most because of their strong links with industry.

"We certainly don't shy away from competition", says Max Tilzer, director of the Alfred-Wegener-Institut for Polar and Marine Research in Bremerhaven. "But to make sure we get a fair chance, industrial and basic research projects have to be treated equally."

Treusch, who is director of both the Helmholtz Society and the Forschungszentrum Jülich, dismisses the fear that, in accordance with prevailing political opinion, basic research could fall behind. "HGF centres of excellence should have hopes instead of fears," he says.

Quirin Schiermeier

South African museums' status 'at risk'

[CAPE TOWN] Biologists at two South African museums are disturbed at government plans to downgrade their institutions from national to provincial status, which they fear could threaten the survival of their collections.

Earlier this year, the Department of Arts, Culture, Science and Technology decided to implement proposals by a 1995 task force to group most of the country's 18 national museums into two institutions from April 1998. This would create two national museums, one in each of the twin capitals of Pretoria and Cape Town.

But some institutions will not be included. In particular, both Natal Museum and the National Museum in Bloemfontein will be downgraded to provincial status, and will in future be funded by their respective provincial governments. The KwaZulu-Natal government, which will be responsible for funding Natal Museum, cut its museum service operating budget this year by more than 51 per cent.

Both museums hold collections of scientific interest. The Natal Museum's mollusc collection, in particular, is by far the largest both on the African continent and on the entire Indian Ocean rim. George Branch, professor of zoology at the University of Cape Town, says: "There is no doubt that this is a national resource which all marine ecologists in the country use for reference purposes. If its funding is reduced, then we risk losing both the collection itself, which requires maintenance, and the taxonomic expertise of those who work on it."

Two researchers at the Natal Museum, Dick Kilburn and Dai Herbert, have



Natal's museum: mecca for molluscs.

requested colleagues worldwide to appeal against the decision, which they claim has been taken without any assessment of the significance of the museum's collections.

In fact, a review committee set up by the Department of Arts, Culture, Science and Technology last year to review the status of the country's museums did not assess the geographical representativeness of the collections at either of the two museums whose status is to be downgraded (see *Nature* 384, 15; 1996). Nor did it make such an assessment of the two natural history museums that will be incorporated into the new national museums, the South African Museum in Cape Town and the Transvaal Museum in Pretoria.

The fate of the fifth natural history museum with national status, the J. L. B. Smith Institute of Ichthyology in Grahamstown, appears undecided. Its collection of 750,000 fish specimens is widely regarded as being of national importance.

Michael Cherry

Nuclear love affair loses its charm

Declan Butler

The strong showing of environmentalists in last month's elections in France has marked a turning point in the history of its nuclear power programme, and raised doubts over the future of French reprocessing.

lionel Jospin, France's Socialist Prime Minister, has thrown into question the overall strategy of the world's foremost nuclear power programme with his confirmation last week that the government intends to "abandon" Superphénix, the only commercial-scale fast-breeder nuclear reactor.

Also at stake is the future of the world's largest reprocessing plant for nuclear fuel, at La Hague on France's northern coast.

Superphénix was built in the 1970s by NERSA, a consortium of French, German and Italian utilities, at Creys-Malville near Lyons. But the 1,250-MW reactor, whose total cost is put at FF60 billion (US\$10.3 billion), has been plagued by technical prob-

lems. It has operated for a total of only a few months since it was opened in 1986.

Jospin's announcement of its closure, made during his maiden speech to the National Assembly last week, represents a major victory for the 'green' movement. Dominique Voynet, the leader of the Green party, Les Verts, and environment minister in the new government, had demanded the closure as part of a pre-election pact with the Socialist party. Les Verts entered the assembly for the first time after the general election earlier this month, when it won eight seats.

Voynet's ministry will be strengthened by the fact that it succeeded in fighting off a last-minute bid to save the reactor by the

nuclear lobby and the Socialist party's Communist allies, who were supported by reactor staff and local politicians. The decision is a serious defeat for those who have until now steered the country's nuclear power programme, free of any significant public scrutiny.

France produces 77 per cent of its electricity from nuclear power, and few expect the country to follow Germany and other countries in phasing out nuclear power. But the shift in the balance of forces between the nuclear lobby and environmentalists suggests that greater public scrutiny is inevitable.

Jospin said that, although France's nuclear industry is important, it should not be "exempt from democratic rules" and that greater independent control is needed.

Following the decision to close Superphénix, many observers in Paris are now questioning the fate of the La Hague plant. Reprocessing plants were originally built to provide plutonium for weapons, and later for a planned park of fast-breeder reactors



Voynet: invoked pre-election pact.

CELIA ERKUL, GAMMA

Japan set to face a lonely future in fast-breeder research

[TOKYO] The Japanese government is determined to press on with the development of fast-breeder reactors, despite France's decision to close Superphénix and calls for the "dissolution" of the organization that operates Japan's Monju fast-breeder (pictured right) following a series of accidents.

At its meeting last Friday (20 June), the Atomic Energy Commission reaffirmed Japan's policy of developing a nuclear fuel cycle that will use plutonium in conventional and fast-breeder reactors. The policy had been put in doubt following a series of accidents at facilities operated by the Power Reactor and Nuclear Fuel Corporation (PNC), including a sodium leak at Monju in 1995 and an explosion at a low-level waste facility last March (see *Nature* 386, 746; 1997).

Some feel that the French decision, announced the day before the commission's meeting, could isolate Japan from other industrialized nations in its nuclear power strategy. But the head of the commission, Riichiro Chikaoka, who also heads the government's Science and Technology Agency (STA), which oversees PNC, declared that "the importance of our nuclear fuel cycle project has not changed [as a result of the French policy]."

The commission's decision came only a few days after the head of a committee set



up by the agency to look into reform of the PNC suggested that a revamped corporation should concentrate on research and development of fast-breeder reactors and nuclear waste disposal. Hiroyuki Yoshikawa, head of the committee and former president of Tokyo University, says that a reformed PNC should concentrate on "pragmatic" research and development. Certain activities should be privatized, transferred to other

organizations or terminated, such as uranium surveys overseas and development of the advanced thermal reactor Fugen.

Yoshikawa's draft proposals, which will be finalized next month, are considerably less severe than suggestions by some politicians that PNC should be "dissolved".

A parallel PNC reform committee of the ruling Liberal Democratic party, headed by Hidenao Nakagawa, a former director-general of the STA, is likely to recommend that PNC should be dissolved. But even this committee is expected to propose the continued development of fast-breeder reactors, according to Kaname Ikeda, director-general of STA's nuclear safety bureau.

The STA's budget for Monju will decrease next fiscal year to about ¥10 billion (US\$88 million) from the current ¥13 to ¥14 billion, says Ikeda. But this is because the facility is "sleeping" following the sodium leak.

But France is the leader in fast breeder technology, and the closure of Superphénix is bound to at least cause a rethink in Japan, the only other country with a substantial fast-breeder programme, says Gerhard Heusener, who heads nuclear safety at the Karlsruhe Nuclear Research Centre, and led German participation in French efforts to explore using fast breeders as plutonium incinerators.

David Swinbanks & D.B.

AP KYODO

that would burn plutonium and breed further plutonium fuel.

The plutonium cycle has dominated French nuclear thinking since the 1970s, when it was predicted that uranium would become scarce and that plutonium would become the fuel of choice.

Twenty years later, uranium prices are lower than ever, and plutonium from reprocessed spent fuel and dismantled weapons is generating a large and unwanted stockpile. But reprocessing plants continue to be built and operated, with the La Hague plant producing 15 tonnes of plutonium a year.

The conversion of Superphénix in 1994 from a power plant to a research reactor aimed at incinerating plutonium may have ended the French dream of developing power-generating breeders. But it still gave the plutonium cycle a second lease of life, by keeping open the unlikely option of developing a park of incinerators, and the associated reprocessing plants that would be needed. Although such research may continue at a reduced level, any plans for industrial-scale fast breeders in the near future have now disappeared.

Controversy over the La Hague plant has also been fuelled recently by a study linking it to increased rates of leukaemia in the region, and claims that levels of radioactivity in discharges from the plant were greater than officially stated (see right).

Incineration remains a "valid research angle", however, claims Gerhard Hübner, who heads nuclear safety at the Karlsruhe Nuclear Research Centre in Germany, and leads the small German participation — about DM 5 million (US\$2.9 million) annually — in the French incineration research programme CAPRA.

But he admits that it is "hard to imagine" using fast neutron reactors to burn plutonium on an industrial scale in the near future, given that this would require building one plutonium furnace for every five pressurized-water reactors.

Debate is likely to grow over the coming weeks with the expected completion of a report commissioned by the government on the recycling of plutonium. "The end of the breeder programme puts into question the rest of the [plutonium] system", argues Mylène Schneider, director of WISE, a Paris-based energy consultancy.

Les Verts has called for a halt to new reprocessing contracts, pointing out that reprocessing is by far the most polluting part of the nuclear fuel cycle. Observers in Paris predict a long, hard battle over the future of the La Hague plant, adding that the outcome is difficult to predict. The plant has been a key component in the management of France's spent fuel and its shutdown would provoke a complete rethink of the French nuclear fuel cycle. □

Cogema's 'arrogance' adds to La Hague's problems

[PARIS] Continuing controversy about the level of radioactivity in discharges from the reprocessing plant at La Hague took a political turn last week when France's environment and health ministers rallied to the support of the environmentalist group Greenpeace.

This followed an incident in which divers from Cogema, the company that operates the plant, were alleged to have removed underwater monitoring equipment that had been installed by Greenpeace.

Greenpeace immediately filed charges against 'X', a standard legal procedure in France, for "theft by an organized gang". Cogema admitted having "confiscated a foreign body". Dominique Voynet, the environment minister, criticized the removal of the equipment, saying "it is not unusual for an independent organization such as Greenpeace to exercise its role of vigilance by carrying out measurements of the discharges".

In a blistering leading article, the newspaper *Le Monde* described Cogema's action as "pitiful". Drawing a comparison with the bombing by the French secret services of the Greenpeace ship *Rainbow Warrior* in Auckland harbour in 1985, the newspaper argued that the latest incident betrayed the persistence of an obsolete "arrogant" belief within the nuclear industry that nuclear activities are none of the public's business.

The pipe that discharges waste from the La Hague plant into the English Channel has become the flashpoint of a lengthy battle between Greenpeace and Cogema. The latest measurements by Greenpeace show that levels of radioactivity in sediments in the vicinity of the pipe exceed European norms.

Officials at the health ministry expressed disbelief that official monitoring could have missed what Greenpeace detected. The government body responsible for carrying out spot checks on Cogema's monitoring, the Office de Protection contre les Rayonnements Ionisants (OPRI), limply explained that it had restricted measurements to the beaches, and had ignored the area around the end of the pipe.

The revelations follow a controversy earlier this year, when exceptionally low tides exposed the mouth of the pipe on a beach open to the public. Measurements made then by CRIIRAD, a respected private nuclear monitoring group, at the request of Greenpeace showed levels of 300 microsieverts per hour at the mouth of the pipe. "A person remaining in the vicinity of the pipe for four hours would have received a dose of radioactivity exceeding the annual maximum dose", says Bruno Chareyron, a CRIIRAD official.

CRIIRAD has protested about the scale of



Channel vision: La Hague's waste discharges have become a focus of attention.

the authorized outputs, claiming that these amount to a figure greater than the discharges of all the world's nuclear reactors taken together. The plant is authorized to release annually 37,000TBq of tritium alone.

CRIIRAD complains that the methods used to determine acceptable levels are available only to Cogema and the government, but not to the public. Chareyron points out that Cogema gathers data on only 4r radionuclides, whereas the UK authorities responsible for monitoring discharges from the Sellafield reprocessing plant look at 14.

CRIIRAD has also found the highly toxic radionuclide iodine-129 in moss within a 7-km radius of the plant, suggesting that the plant was also producing air pollution. Cogema and OPRI had not detected this, Chareyron claims.

Public concern about such claims has been heightened by preliminary research results suggesting that rates of childhood leukaemia in the region are higher than expected.

Following the success of Les Verts in last month's general election, the political tide in France now seems to be turning in favour of environmental organizations. Voynet said last week that she intends to demand an independent study of the levels of radioactivity being dumped from the plant into the channel, as well as a public inquiry into the levels of authorized outputs. And Bernard Kouchner, the junior minister for health, instructed OPRI to carry out a detailed radiological survey of the area around the pipe.

A critical test of the new government's position will come this autumn, when 15 European countries are scheduled to sign the Convention for the Prevention of Marine Pollution, which includes the ultimate goal of reducing radioactive emissions from reprocessing plants to close to zero.

D.B.

Academy backs calls for caution on gene sequence patents

[WASHINGTON] The council of the US National Academy of Sciences has joined the National Institutes of Health (NIH) and the Human Genome Organization in protesting at reports that the US Patent and Trademark Office has agreed to issue patents on expressed sequence tags known as ESTs (see *Nature* 385, 670; 1997).

Bruce Alberts, president of the academy, last week wrote to Bruce Lehman, the Commissioner of Patents and Trademarks, expressing the fear that broad patents on ESTs "could be devastating to the study of human biology and disease by limiting access to the finite set of human genes". Those holding such patents could claim the rights to control all potential uses of a particular gene sequence.

Lehman had earlier written to Harold Varmus, the director of NIH, saying that ESTs might be patentable "under appropriate and limited circumstances". Writing on behalf of the academy's council, Alberts says it is "crucial" to define clearly how this phrase is interpreted, and how the appropriate scope of claims for ESTs will be judged. "It would be sad indeed if patent policies diminished the pace of discovery or wealth of practical applications," says Alberts. "Yet patents that allow an early group of inventors who have disclosed little new knowledge to constrain the actions of subsequent investigators threaten to do just that."

Medical technologies lead rise in patent bids

[MUNICH] The number of patent applications filed at the European Patent Office in Munich rose by 9.2 per cent in 1996. Medical and veterinary technologies headed the list in terms of numbers, with more than 5,400 patent applications. Their 10.2 per cent growth compared with 1995 is mainly due to the increased number of applications for US discoveries in these two fields.

Electronic communication technologies (up 13.9 per cent) and vehicle technology (up 23.4 per cent) showed the highest rates of increase. In contrast, overall patent applications in organic chemistry rose by only 3.3 per cent, and in organic macromolecular compounds applications fell by 1.1 per cent.

Anthropologists win science book prize

[LONDON] This year's £10,000 (US\$16,500) UK Rhône-Poulenc Science Book Prize has been won by a husband-and-wife team of palaeoanthropologists for a book about the

discovery of the most complete skeleton of *Homo erectus* ever found.

Alan Walker, professor of anthropology and biology at Pennsylvania State University, shares the award with the science writer Pat Shipman, who is an adjunct professor of anthropology at the university, for *The Wisdom of Bones: In Search of Human Origins*. The book, which is published by Knopf/Weidenfeld, tells the story of Walker's discovery in 1984 of 'Nariokotome Boy', a 1.5-million-year-old *Homo erectus* skeleton found in Kenya.

Health 'needs higher profile' in Euro research

[MUNICH] The European Commission's proposed fifth Framework research programme should be revised to include a thematic programme dedicated to health research, according to a new report from the European Science Foundation's standing committee for the medical sciences.

The committee, which represents more than 30 European biomedical research funding agencies, says that health issues could get lost within the proposed programme, called "Unlocking the Resources of the Living World and Ecosystem", because of its breadth and heterogeneity. To capitalize fully on Europe's biomedical research potential, it argues, health requires its own programme, with clearly defined key actions such as research into the ageing population.

Radiation accident kills Russian researcher

[MOSCOW] Alexander Zakaharov, a Russian nuclear researcher, has died after being taken to hospital suffering from radiation exposure after an accident on 17 June at the Arzamas-16 nuclear research centre, 200 miles east of Moscow. Zakaharov is believed to be Russia's first nuclear accident victim since the Chernobyl disaster in 1986.

But the Russian government claims to have found no evidence that radioactive material leaked outside the centre's 'site number 8' where the accident is understood to have taken place. Officials say that radioactivity remains at background levels.

Grants to support young German scientists

[MUNICH] Highly qualified young German scientists with academic ambitions but lacking sufficient experience for a professorship have been targeted for support by the Max Planck Society. At its annual meeting in Bremen two weeks ago, the society approved the establishment of a new funding programme to employ talented scientists who have won their *Habilitation* — the German teaching-and-research qualification

required for a university appointment with the Max Planck Society — but who have not yet found a permanent position. The stipends, each worth around DM100,000 (US\$57,500) a year, will be awarded for a maximum of three years.

Sweden 'will not quit molecular biology lab'

[MUNICH] Sweden will not after all withdraw from membership of the European Molecular Biological Laboratory in order to save money, according to a statement by Carl Tham, the country's research minister. This statement was published earlier this month to allay fears that such a step could be necessary as part of a package to save SKr150 million (US\$19.5 million) from next year's research budget (see *Nature* 387, 224; 1997). "The general cuts in research spending that have been decided upon [by the government] will be implemented in some other way, and this will be reported in the autumn's budget bill," said Tham.

Blood lab boss jailed for neglect over HIV test

[MUNICH] A former official of Haemoplas, a German blood products company, has been sentenced to prison for six-and-a-half years for neglecting to test blood samples sufficiently for human immunodeficiency virus (HIV). Günter Eckert was the head of the company's blood testing laboratory. The court accepted that he had caused the death of a woman who died of AIDS from a Haemoplas blood product infected with HIV (see *Nature* 376, 628; 1995). Eckert was convicted of lethal injury. The prosecution had called for life imprisonment.

Washington State gets 'shock physics' grant

[WASHINGTON] The US Department of Energy announced plans on Monday to award Washington State University \$10 million over the next five years to perform fundamental research in shock physics for its nuclear weapons program. The research, which will be unclassified and involve no nuclear materials, will support the agency's efforts to maintain the safety and reliability of US nuclear weapons in the absence of testing.

According to Victor Reis, assistant secretary for defense programs, the DOE funding will create an Institute for Shock Physics at the university, which will explore the very rapid compression of materials, shock-induced chemical changes, detonation science and the dynamic response of materials at large compressions and deformations. Such work will lead to improved understanding of the effects on weapons materials and components.

Serious flaws in UK funding system

Sir — The article by Wennerås and Wold (*Nature* **387**, 341–343; 1997), containing clear proof of nepotism in a funding body, is fascinating.

I have had many discussions with contemporaries and colleagues about the way in which decisions are made by the major UK funding bodies, and there is general agreement that the system is seriously flawed.

Everyone has stories about autocratic committee chairmen with whom no one dares to disagree, referees with a study competitive to that of the applicant, committee members who take up ideas for their own groups a year or two after the originator's application was turned down, and committees where more than half the grants under discussion were awarded to members of the committee.

More recently, in Europe, I heard of two European Commission contracts awarded to a group that had been removed from its own national register because of poor performance and no publications. Of course, no one has had, until recently (see D. F. Horrobin, *The Lancet* **348**, 1293–1295; 1996), any good idea how the system might be changed to avoid lack of objectivity, bias and the unfair apportioning of huge sums of money, often repeatedly.

I have been involved in making grant applications to the major funding bodies since the mid-1960s. In all that time, it was until recently never necessary to object to fair or even relatively unfair criticism. However, I wrote last year to one of the major funding bodies about the careless way in which our application was both refereed and subsequently dealt with by the committee. It was not prompted by sour grapes but by the indisputably incorrect

statements made in their letter of rejection.

My letter was designed, first, to explain how much more helpful it would be to those of us 'on the other side' to receive feedback before the committee meeting that decides our scientific futures. Grant applications are dealt with in this way by at least one major UK funding body. Secondly, I wished to point out the extent to which the outcome can be substantially affected by luck and by objectivity (or lack of it) from both referees and committee members.

The unsatisfactory reply that I received contained the comment: "You are free, of course, to submit a fresh application if that is substantially different from this one." Given that the reasons for turning down the application could not be discussed, are we to play some sort of guessing game as to how our "important study" (the committee's words) should be redesigned? If I am to go by the referees' reports, then very little needs to be altered. Am I therefore to take a stab at suggesting something completely different which is trendy and of academic interest but may not improve the welfare of patients to the extent that I believe our present protocol may do?

This is a difficult and dispiriting situation, which erodes research time, thinking time and confidence. It is not surprising that nothing will induce many graduate students (male or female) to take up scientific research as a career these days.

Linda M. Castell

Cellular Nutrition Research Group,
University Department of Biochemistry,
South Parks Road,
Oxford OX1 3QU, UK
e-mail: cat@bioch.ox.ac.uk

Sir — The research that Wennerås and Wold have been able to do and therefore the issues they raise are crucially dependent on the Swedish commitment to freedom of information. The findings, however, are of relevance to the global research community.

In the British case, our pathological love of secrecy means that it is simply impossible to monitor equity of treatment, either in the procedures for the allocation of grants or in the recognition of scientific excellence. Bodies from the research councils to the Royal Society should find this study food for thought, as they, like the judiciary, are institutions where pale males have been for too long chosen to overrepresent themselves for the health of either science or society.

For that matter, *Nature's* innovatory footnote drawing attention to the gender of the three referees was itself both smug and sexist. A more worthy reaction would have been for *Nature* to concentrate first on the possibility of a beam in its own eye. In the same issue, your account of the "Science Wars" (*Nature* **387**, 331–335) manages to erase the fact that Gross and Levitt had attacked, most vitriolically and extensively, the feminist critics of science. Indeed, as I recall their book, the single most attacked person was the distinguished US philosopher of science Sandra Harding. By focusing exclusively on the fights between the boys, *Nature* managed to make even the "Science Wars" another pale male story.

Is it any wonder that some of us think that the tide is rather unlikely to turn?

Hilary Rose

4 Lloyd Square,
London WC1X 9BA, UK
e-mail: h.rose@open.ac.uk

Cannibalism and kuru

Sir — In his review of *Deadly Feasts: Tracking the Secrets of a Terrifying New Plague* (*Nature* **386**, 565; 1997), Robert Desowitz repeats a fundamental error with regard to the aetiology of kuru. He writes: "Gajdusek and his Australian colleagues, Michael Alpers and Vincent Zigas, showed that cannibalism was the cause of kuru. The Fore women and children, but not the men, ate — totally consumed — their dead relatives."

That cannibalism might be the route to kuru transmission was discovered by two anthropologists, Shirley Lindenbaum and Robert Glasse, five years after Gajdusek first arrived in the Eastern Highlands of Papua New Guinea, years in which every attempt to unravel the cause of kuru had failed. In

1961, they spent nine months with the Fore searching for a possible genetic cause for kuru. None existed.

Their second visit proved more fruitful. A New Zealand neurologist and epidemiologist, Dr R.W. Hornbrook, suggested that they try to answer the question "what is it that the adult women and the children of both sexes in the Fore tribe are doing that the adult men are not?"

They found that the women, who prepared the bodies for burial, occasionally ate part of the flesh and of the steamed brain tissue. But far from their "totally consuming their dead relatives", or cannibalism being a ritual in the Fore tribe, it was a casual practice that had infiltrated from tribes in the south of the region within living memory. The Fore themselves said that "it was after the first aeroplane flew over that

we tried cannibalism for the first time".

Back at the US National Institutes of Health, Gajdusek and Joe Gibbs now tried to induce kuru in nonhuman primates by feeding them with kuru-infected human brain. They failed to induce the disease, even when they inserted the infected tissue directly into the stomach through a gastric tube. So they believe that direct inoculation into the bloodstream or through the mucous membrane was the route of infection. Women and children with cuts and sores on their hands would frequently rub their eyes and noses when handling dead bodies, so any infectious particles could enter the body quite easily.

June Goodfield

International Health & Biomedicine,
The Manor House, Alfriston,
East Sussex BN26 5SY, UK

Tories out on cue

Sir — In the issue of 30 September 1993, Landsberg, Dewynne and Please¹ used a formula of mine² to predict how long the Conservative government in Britain would continue its rule. My formula, based on the Copernican principle, stated that, if your location is not special, there is a 95 per cent chance that the future observed longevity of whatever you are observing now will be at least 1/39th as long as its past observed longevity but less than 39 times as long as its past observed longevity. Since the Conservative party had been in power for 14 years in 1993, Landsberg *et al.* estimated with 95 per cent confidence that it would remain in power for at least 4.3 more months but less than 546 more years.

The Conservative party lost power 3.6 years later, on 2 May 1997, in agreement with the prediction. This was a reasonable application of the formula, because it was prompted by the publication of my paper, whose publication date had no relation to British politics (unlike this letter). Indeed the claim of my paper is that its date of publication, 27 May 1993, is not special (except relative to itself, of course). The Copernican hypothesis can be tested: take an ensemble of things observable on my paper's date of publication and ask

whether their future observability is correctly predicted by the formula 95 per cent of the time (see ref. 3 for some examples, from Broadway and off-Broadway plays and musicals to world leaders' terms of office).

There have been other successful predictions. My paper used the above (95 per cent confidence) formula explicitly to predict that the human spaceflight programme would last at least 32/39ths of a year more — it has — and that *Nature* would continue publication for at least 3.15 more years — it has. The upper limits for these, 1,248 years and 4,800 years respectively, can be checked in the future. If the formula works well in all these applications, we may well take seriously my paper's (95 per cent confidence) prediction for the (200,000-year-old) human race as well: that its future longevity will be at least 5,100 years but less than 7.8 million years.

J. Richard Gott III

Department of Astrophysical Sciences,
Princeton University,
Princeton,
New Jersey 08544, USA
e-mail: jrg@astro.princeton.edu

1. Landsberg, P. T., Dewynne, J. N. & Please, C. P. *Nature* **365**, 384 (1993).

2. Gott, J. R. *Nature* **363**, 315–319 (1993).

3. Gott, J. R. in *Astronomical Society of the Pacific Conference Series* (eds Trimble, V. & Reisenegger, A.) **88**, 140 (San Francisco, 1996).

Stranglehold on science

Sir — The attempt by P.-L. Chau and W. Y. Chau (*Nature* **386**, 754; 1997) to refute Audrey Wells' observation of the dearth of individuality in Asian cultures as a consequence of Confucianism (*Nature* **386**, 14; 1997) fails completely. They question whether it is "statistically significant to generalize from Western musical education in Korea and Japan in the past 30 years or so to arrive at a general conclusion about a complex 2,500-year-old philosophy that has deeply influenced the whole of East Asia". They then proceed to produce the same generalizations for which they so eagerly criticized Wells, by attempting to extrapolate present-day Asian culture from traditional Confucian philosophy.

Lacking any understanding of Joseph Needham and his writings, Chau and Chau question whether dualism or "paradoxes" have ever been connected to Taoism. Duality and paradox are intimately connected with Taoist philosophy, as any casual perusal of early Taoist texts would confirm. Furthermore, in *Science and Civilization in China*, Needham raises the point clearly and succinctly that, in fact, one of the possible reasons for the lack of a scientific revolution in China was the possibility that Confucianism imposed a

collective stranglehold on Chinese intellectual advancement. In addition, one critical point that is studiously not addressed by Chau and Chau is Needham's observations regarding the culturally defined roles of science in Chinese society, in which practitioners of science (in the broadest sense) traditionally served the authoritarian needs of the state (royal court), particularly in close conjunction with well-defined élite classes that promulgated the systematic abuse of the predominantly submissive agrarian population.

To confound matters further, Chau and Chau seem not to understand that the concept of individualism has radically different meanings within the broader contexts of Asian and Western societies. They are also incorrect in stating that authoritarianism was unknown in the original writings of Confucian philosophers. On the contrary, Confucianism developed out of a society with a highly regimented social structure with little internal mobility; thus, virtue as a putatively beneficial quality was, and is, traditionally imposed from above in Chinese society. To the extent that moral courage existed in Chinese society, particularly during the Tang and Qing dynasties, it was a consequence of individual behaviour (and varied from

person to person), not Confucianism. It was certainly not an idealistic consequence of well-travelled paths having been trodden in the search for greater meanings. Filial piety, therefore, is necessary for maintaining social harmony in Asian society. Confucian philosophers and scholars were hardly the only ones to have understood this point. Mao reformed this concept, and then used it to great ill-effect in his egotistical attempt to dominate Chinese society. Mao understood the Confucian Analects as well as any Chinese — probably more than most — and evidently to a much greater extent than the authors of the recent correspondence.

Peter Cohen

*Department of Chemistry,
Columbia University,
MC 3150, Havemeyer Hall,
3000 Broadway,
New York, New York 10027, USA
e-mail: cohen@chem.columbia.edu*

You read it here first

Sir — S. H. Friedman and J. O. M. Karlsson recently reported what they believed to be a novel paradigm (*Nature* **385**, 480; 1997).

They presented a graph of the proportion of articles in the Medline database with the word 'novel' in the title or

abstract, plotted as a function of the year of publication. They extrapolated, from the exponential rise in the use of the word, that by 2020 all scientific papers will claim novel findings. It is ironic that their observation is itself not novel.

Nature published correspondence in 1991 (**350**, 9; 1991) in which we noted the accelerating use of the word 'novel' and documented it with a graph of the frequency of use of 'novel' and two control words ('control' and 'unusual') in Medline database records. In addition, we recommended that authors reserve 'novel' for strikingly new discoveries, lest the word lose its impact.

We are sorry to see, from the updated data provided by Friedman and Karlsson, that our recommendation has had no discernible impact. It appears that the editors of *Nature* recognize the dire consequences of the novelty explosion and will provide regular updates of the situation until the message takes hold.

Robert W. Levis

*Fred Hutchinson Cancer Research Center,
Seattle,
Washington 98109, USA*

Steven Henikoff

*Fred Hutchinson Cancer Research Center,
and Howard Hughes Medical Institute,
Seattle,
Washington 98109, USA*

Europe ambivalent on biotechnology

Throughout Europe, there is widespread lack of trust in the ability of governments and other public authorities to deal effectively with people's concern about biotechnology applications.

Biotechnology and the European Public Concerted Action group*

Many Europeans are uneasy about modern biotechnology, particularly about new genetic technologies. Although there is widespread support for 'traditional' medical applications in the fields of diagnosis and treatment, few approve of the use of transgenic animals for research or for applications such as transplantation of organs into humans (Fig. 1). There is also a striking mismatch between the traditional concern of regulators with issues of risk and safety, and that of the public, which centres on questions of moral acceptability. These are some of the conclusions to emerge from the latest Eurobarometer survey, designed to find out what people think about biotechnology (see box on following page). The main lesson of the survey is that public confidence in emerging applications of biotechnology cannot be taken for granted.

Conventional wisdom holds that knowledge is a crucially important determinant of support for science and technology — the more informed the public, the more likely it is to be supportive (see, for example, ref. 1). But comparison of the new (1996) Eurobarometer survey with earlier ones in 1993 (ref. 2) and 1991 indicates that although the public's knowledge of relevant basic biology has increased slightly, optimism about the contribution of biotechnology and genetic engineering to improving our way of life has actually declined. Furthermore, the new survey shows that knowledge is poorly correlated with support for all the applications described in Fig. 1.

Thus, as already discovered by other industries trying to introduce controversial technologies (such as the nuclear industry), more knowledge does not necessarily lead to greater public acceptance. But the situation with respect to biotechnology is more complex. The new survey suggests, for example, that people with greater knowledge are more likely to express a definite opinion about biotechnology; but this opinion can be positive or negative. It has been claimed that 'mismatches' between scientific and lay assessments of risk are responsible for public resistance to new technologies³. Public debates about some aspects of biotech-

nology have been dominated by the issue of risk, whereas in others moral considerations have been more prominent. Figure 1 shows that people see all biotechnology applications as potentially useful, but those involving crop plants, food production, the use of transgenic animals for research and xenotransplantation (trans-species organ transplants) are seen to involve risks; whereas only the use of transgenic animals for research and xenotransplantation are thought of as morally unacceptable. At first sight, this pattern implies that use, risk and moral acceptability are all likely to be strongly correlated with overall levels of support for specific areas of biotechnology.

Surprisingly, however, multiple-regression analysis indicates that although moral acceptability and use are strong predictors of support as measured by encouragement (moral acceptability, average $B = 0.54$; use, average $B = 0.35$, where B is an index of the strength of association), risk has very low predictive value (average $B = 0.04$). Only in

the case of food production does risk perception appear to be more than a trivially small predictor of encouragement. The pattern of results across the six applications in Fig. 1 suggests that perceptions of usefulness, riskiness and moral acceptability could be combined to shape overall support in the following way. First, usefulness is a precondition of support; second, people seem prepared to accept some risk as long as there is a perception of usefulness and no moral concern; but third, and crucially, moral doubts act as a veto irrespective of people's views on use and risk.

The finding that risk is less significant than moral acceptability in shaping public perceptions of biotechnology holds true in each EU country and across all six specific applications described in Fig. 1. This has important implications for policy-making. In general, policy debates about biotechnology have been couched in terms of potential risks to the environment and/or human health. If, however, people are more swayed

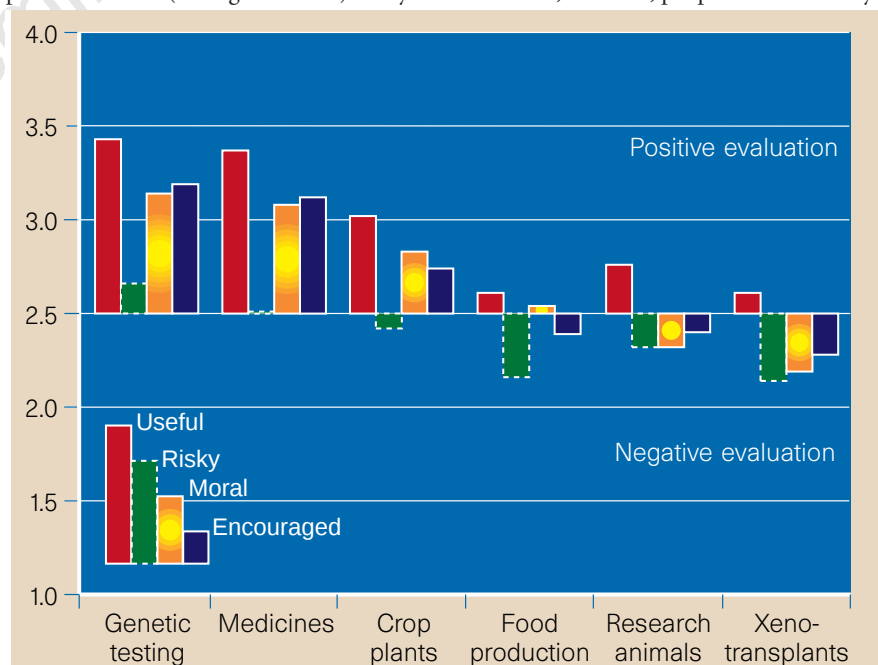


Figure 1 Perceived use, risk and moral acceptability as determinants of public support. Respondents were asked whether they thought each of six biotechnologies was useful (red), risky (green), morally acceptable (yellow) and whether it should be encouraged (blue). Mean scores across Europe are given on a four-point scale, in which 2.5 is the 'neutral' point. Genetic testing, using genetic tests to detect inheritable diseases such as cystic fibrosis. Medicines, introducing human genes into bacteria to produce medicines or vaccines, for example to produce insulin for diabetics. Crop plants, transferring genes from plant species into crop plants to increase resistance to insect pests. Food production, using modern biotechnology in the production of foods, for example to make them higher in protein, keep longer or change in taste. Research animals, developing genetically modified animals for laboratory research studies, such as a mouse with genes that cause it to develop cancer. Xenotransplants, introducing human genes into animals to produce organs for human transplants, such as into pigs for heart transplants into humans.

*This article has been written by an international team of researchers working as part of a Concerted Action of the European Commission (B104-CT95-0043) administered on behalf of Directorate General XII by Andreas Klepsch. For details see box overleaf. Address for correspondence: G. Gaskell, Department of Social Psychology, London School of Economics, Houghton Street, London WC2A 2AE, UK (e-mail: gaskell@lse.ac.uk).

Eurobarometer survey

The Eurobarometer on Biotechnology (46.1) was conducted during October and November 1996. The survey conducted in each EU (European Union) country used a multi-stage random sampling procedure and provided a statistically representative sample of national residents aged 15 and over. The total sample within the EU was 16,246 respondents (about 1,000 per EU country). The survey questionnaire was designed by the authors as part of a larger study involving the comparative analysis of public perceptions, media coverage and public policy in relation to biotechnology from 1973 to the present.

Members of the Concerted Action who contributed to this article are: W. Wagner & H. Torgerson, Austria; E. Einsiedel, Canada; E. Jelsoe, H. Fredrickson & J. Lassen, Denmark; T. Rusanen, Finland; D. Boy & S. de Cheveigne, France; J. Hampel, Germany; A. Stathopoulou, Greece; A. Allansdottir, Italy; C. Midden, Netherlands; T. Nielsen, Norway; A. Przystalski & T. Twardowski, Poland; B. Fjaestad, S. Olsson & A. Olofsson, Sweden; G. Gaskell, J. Durant, M. Bauer & M. Liakopoulos, UK.

The project is coordinated by M. Bauer, J. Durant and G. Gaskell.

by moral considerations, public concern is unlikely to be alleviated by technically based reassurances and/or regulatory initiatives that deal exclusively with the avoidance of harm.

It is because of the issues of risk and safety that modern biotechnology is so extensively regulated in Europe. In the new survey, people were asked which bodies they thought best placed to regulate modern biotechnology (Fig. 2). On average, more Europeans preferred international organizations such as the United Nations and the World Health Organisation to either their own national or pan-European public bodies. Self-regulation by scientific organizations also rated highly. These results confirm the trend, observed by many others, of an increasing lack of confidence in national political institutions. They also show, however, that biotechnology is seen as having transnational consequences which national bodies are powerless to influence.

The survey highlights public concern that may have arisen as a result of lack of trust. For example, 74% of respondents consider that genetically modified food should be labelled; 60% believe that there should be public consultation about new developments in biotechnology; 53% say that current regulations are insufficient to protect people from the risks of biotechnology; and 39% think that religious authorities should be involved



Figure 2 Who should regulate biotechnology? People were asked which of a selection of bodies they thought best placed to undertake this function.

in the regulation of biotechnology. If biotechnologists and industry regulators are to command public confidence, they must incorporate openness and wide consultation into the policy-making process.

Do people discriminate among information sources for different issues? We asked people to select from a list of 12 institutions the one they thought most likely to be honest about two areas of modern biotechnology

(Fig. 3). For new genetically modified food crops, people had most trust in environmental organizations, whereas for xenotransplantation people preferred the medical profession. Thus, people do discriminate among sources of information, and trust is strongly issue-specific. The medical profession remains one of the most widely trusted institutions within its own sphere of competence, as to a lesser extent do environ-

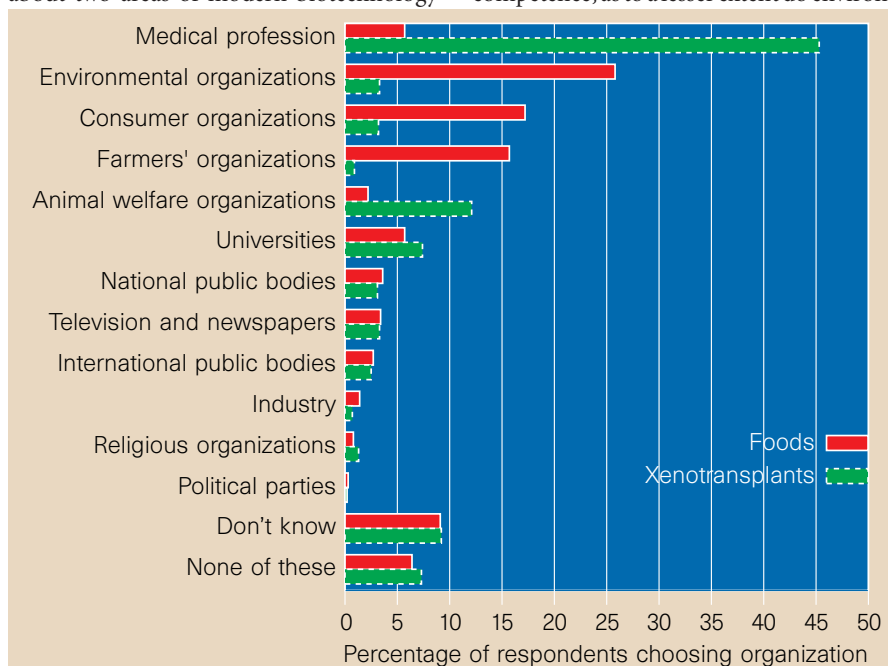


Figure 3 Who can be trusted to tell the truth about biotechnology? People were asked in which of several organizations they had confidence to tell the truth about new genetically modified food crops grown in fields (red) and the introduction of human genes into animals to produce organs for human transplants (green).

mental organizations in theirs. National political institutions fare badly in public estimation.

In an increasingly complex world, it has been said that trust is a functional substitute for knowledge. Particularly in situations of high uncertainty, lack of trust could become an important determinant of the way issues are viewed: in the absence of trust, perceived risks and moral dangers proliferate and appear greater. For all three main groups of biotechnologies discussed here (medical, agricultural/food, and animal experiments), people who express trust in public authorities tend also to have a systematically more positive view: they are more likely to say that biotechnology should be encouraged; to regard it as morally acceptable; and to view it as less risky. The largest effect of trust is found in the area of agricultural and food biotechnologies, which is of course highly relevant to current public debates on bovine spongiform encephalopathy (BSE) and genetically modified foods.

Thus far, we have treated Europe as if it were a single entity, which of course it is not. In Austria, for example, the opposite of the trend described in the paragraph above applies: those who express trust in public authorities tend also to express opposition to agricultural and food biotechnologies. Indeed, the Austrian government is explicitly opposed to the introduction of genetically modified foods such as soya and maize into the European market⁴. The countries where support for the biotechnology applications described here is greatest are Portugal and Spain, followed by Belgium, Finland and Greece (shown in red in Table 1), whereas those where support is least are Austria and Germany, followed by Denmark, Sweden and Luxembourg (shown in blue).

Some national characteristics associated with support and opposition are shown in Table 1. In general (and with the significant exception of Finland), we find that the public in supportive countries tends to have low levels of contact, low levels of knowledge, a 'menacing' image (see Table 1 footnote for explanation) and many expectations. There is also a relatively relaxed public attitude to risk and regulation. This is a pattern that we might expect to find in less industrialized countries that do not possess a well-developed biotechnology industry and where there is not much public participation in debate about biotechnology.

By contrast, the public in countries where most people oppose biotechnology tends to have high levels of contact, high knowledge, a matter-of-fact image and low-to-moderate expectations. At the same time, there tends to be public anxiety about risk and regulation. This is a pattern we might expect in more industrialized countries that possess a well-developed biotechnology industry and a relatively high level of public participation

Table 1 National attitudes to biotechnology

Attitudes to applications of biotechnology				
Low encouragement Transgenic animals	A D DK Lux S UK IRL NL I B FIN	F	GR E P	High encouragement Transgenic animals
Low encouragement Medical		A, D	DK Lux S IRL NL UK I F B FIN GR E P	High encouragement Medical
Low encouragement Agricultural and food	A D DK Lux S	F NL UK IRL I GR	B FIN E P	High encouragement Agricultural and food
National characteristics				
High contact	A D DK Lux S FIN	F I NL UK B	IRL GR E P	Low contact
High knowledge	DK S NL UK FIN	D Lux F I B	A IRL GR E P	Low knowledge
Matter-of-fact image	DK S NL UK FIN	Lux F I B E	A D IRL GR P	Menacing image
Low positive expectations	A D NL	Lux S I UK B GR FIN	DK IRL F E P	High positive expectations
Low negative expectations	A S NL FIN	D Lux DK I IRL B	UK F GR E P	High negative expectations
Anxious about risk and regulation	A S DK I F	D Lux B GR P	IRL NL UK FIN E	Relaxed about risk and regulation

Summary of aggregate scores for individual EU countries on a series of indicators in the survey. A, Austria; B, Belgium; D, Germany; DK, Denmark; E, Spain; F, France; FIN, Finland; GR, Greece; I, Italy; IRL, Ireland; Lux, Luxembourg; NL, Netherlands; P, Portugal; S, Sweden; UK, United Kingdom. Red, countries supporting biotechnology; blue, countries opposing it (dark shades indicate stronger support or opposition); black, mixed support and opposition. Top, countries are scored negative, neutral or positive on three main groups of biotechnological applications: transgenic animals (animals for research + animals for xenotransplantation); medical (genetic testing + medicines or vaccines); and agricultural and food (crop plants + food production). Bottom, countries are scored high, medium or low on six national characteristics.

'Contact' is derived from two questions designed to measure whether people had heard or talked about biotechnology before. 'Knowledge' is based on a scale of factual questions about relevant basic biology. 'Image' captures popular imagery associated with biotechnology. Respondents were invited to agree or disagree with propositions such as: "ordinary tomatoes do not contain genes while genetically modified tomatoes do"; and "by eating a genetically modified fruit, a person's genes could become modified". These questions were not primarily intended to capture levels of objective knowledge but to measure the extent to which respondents possessed menacing images of biotechnological products. 'Positive and negative expectations' were based on a series of ten paired questions describing a positive and a negative outcome of biotechnology that could happen in the next 20 years. A positive outcome was "curing most genetic diseases" whereas the negative one was "creating dangerous new diseases". 'Relaxed about risk' means that respondents believe that current regulations are sufficient and agree that some risk must be accepted in the interests of economic competitiveness.

in debate. Once again, however, there are exceptions. Austria, for example, combines high contact with low knowledge and 'menacing' image. (Apart from Austria, the countries where most people oppose biotechnology were among the first to introduce biotechnology regulations.)

Across the European Union, then, it seems that some countries in which biotechnology is best established are among the least supportive, whereas others in which the science and industry are in their infancy are the most supportive. This result is not as paradoxical as it appears: in the former group, familiarity with biotechnology has provided greater opportunity for the emergence of concern; whereas in the latter, the potential economic importance of biotechnology is paramount.

However, it would be naive to pretend that the detailed pattern set out in Table 1 has easy or obvious explanations without a far better understanding of national cultures than surveys alone can provide. In Austria, for example, various factors, including general concern for environmental issues and a relatively recent encounter with biotechnology, coincident with anxiety over membership of the European Union, have conspired

to raise the temperature of public debate about biotechnology almost to boiling point.

For now, it is perhaps enough to point out that these Eurobarometer findings echo the theme of the "risk society" as discussed by writers such as Ulrich Beck and Anthony Giddens^{5,6}. Like the European public in our survey, these authors do not find the language of objective risk assessment adequate, arguing that risks are fundamentally moral and political. Our data suggest that large sections of the European public are deeply ambivalent about much of modern biotechnology. The prevailing focus of this ambivalence appears to be moral, a collection of anxieties about unforeseen dangers that may be involved in a range of technologies that are commonly perceived to be 'unnatural'. □

1. Evans, G. & Durant, J. *Public Understand. Sci.* **4**, 57–74 (1995).
2. *Biotechnology and Genetic Engineering: What Europeans Think About it in 1993* (INRA (Europe), Brussels, 1993).
3. Bauer, M. (ed.) *Resistance to New Technology: Nuclear Power, Information Technology, Biotechnology* (Cambridge Univ. Press, 1995).
4. Abbott, A. *Nature* **386**, 745 (1997).
5. Beck, U. *Risk Society: Towards a New Modernity* (Sage, London, 1992).
6. Giddens, A. *The Consequences of Modernity* (Cambridge Univ. Press, 1990).

Symmetry as destiny — taking a balanced view of IQ

Steve Blinkhorn

Genetic vulnerability during development may be responsible for environmental effects on intelligence. Taken with a study showing that the heritability of IQ does not change with age, could these results mean that nature and nurture are no longer in opposition?

The study of human intelligence has always had its feet in biology and its head in the clouds — the recent encounter between Gary Kasparov and Deeper Blue showed what can be achieved in confronting human intelligence with methods that have no plausible claim to mimic its cognitive processes. Two studies at the other end of the spectrum now serve as a reminder that intelligence operates on a substrate of wetware, not silicon.

In the 6 June issue of *Science*, McClearn *et al.*¹ described a study of octogenarian Swedish twins, in which they found that the heritability of psychometric intelligence remains much the same in old age as in youth. And in this week's *Proceedings of the Royal Society*, Furlow *et al.*² show that there is a modest, but replicated, association between IQ and asymmetry in bilateral human physical traits. These observations lead to nicely paradoxical conclusions, relating environmental influences to the genetic basis of differences in IQ.

Research into heritability and the biological correlates of measured intelligence tends, unfortunately, to descend into a kind of cognitive Calvinism. Heritability estimates take on the mantle of constants of nature, social pessimism holds sway, and all concerned resign themselves to predestination due to unknown genes. Because there has long been evidence for the heritability of IQ, the idea took hold that there must, therefore, be genes for intelligence — just as theories now circulating suggest that there may be genes for homosexuality. Of course, ideas about the heritability of intelligence pre-date the notion of a gene, and they go back certainly as far as Francis Galton.

McClearn *et al.*¹ have found that estimates of the heritability of IQ amongst octogenarians are of much the same magnitude as estimates from younger age-groups (of the order of 0.60). This result is surprising only because common sense suggests that, with the passage of time, the effects of experience will become increasingly important in determining cognitive abilities. So, the influence

of heredity should become proportionately less. But if the Swedish study is to be believed, environmental influences on IQ have their effects early. This is hardly controversial stuff and, indeed, potential factors (such as atmospheric lead, maternal nutrition and fetal alcohol syndrome), which indicate a biological rather than a strictly psychological influence, will always focus on the earliest stages of life.

Unfortunately, having relatively stable estimates of the heritability of IQ advances our knowledge rather little. Because we have no sensible systems by which to measure either the variability of environments or the variability of genetic influences, we are left with a ratio of two unknowns. The only conclusions that are left to be drawn are socio-political — heritability of IQ is so great (or small) that educational or social programmes will have insignificant (or enormous) effects.

The problem is, opinion is split three ways as to how to handle the raw observation that observed intellectual capacity varies widely. Those with their heads in the clouds — artificial-intelligence researchers and constructors of psychometric tests (sky pilots for short) — mark their progress by results. Intelligence is about handling difficult tasks successfully, and doing well when you started off not knowing what to do. These sky pilots aim to keep at least two steps ahead of those who find algorithmic solutions to cognitive tasks, and they provide measures which act as criterion variables in studies of heritability and the biological correlates of intelligence.

The second group — the biological determinists or chiropracists (podiatrists if you must) — have their feet firmly planted in biology. They, at worst, suppose that the most we can do is to observe biological determinism, and then look for simple models to explain why variation in intelligence reflects biological variation. Such models include differences in neural efficiency, speed of conduction, or error rates. But whatever you do, don't confuse them by pointing out

the possibility of mental processes that are not strictly reducible to brain processes.

In between are the colonizers from cognitive psychology, who betray their insecurity by adopting the term 'cognitive scientists'. They propose to explain away — rather than to explain — the brute facts of variability in cognitive capacity. They supply the algorithms for solving problems which keep the sky pilots on their toes. For about 25 years they have been promising to account for variability in IQ in terms of cognitive-processing models. (The first positive results are still eagerly awaited.)

Results, processes, substratum — the three foci of interest that divide those who study human intelligence. But somewhere in there is a mind-brain barrier, a conceptual and philosophical divide that stands in the way of progress. The artificial-intelligence and psychometric researchers may set problems and recognize the results as showing the external hallmarks of intelligence, but they have little to say about how to make people more intelligent, or overcome handicaps to achieve their potential. Cognitive scientists have shown how styles of problem that were once thought to require poorly understood human capacities can be reduced to computational algorithms. Yet they have contributed essentially nothing to our knowledge of how some people, when presented with a mass of confusing data, can clearly see the underlying structure of a new problem, whereas others cannot. And the biological determinists can point to over 100 years of consistent findings, but not a lot of hope on offer.

This is where the paper by Furlow *et al.*² seems to break new ground. They have studied fluctuating asymmetry — an asymmetry of usually symmetrical bilateral traits, which occurs as a result of biological stress during development. Fluctuating asymmetry is assessed by an index of variability in the dimensions of feet, fingers, wrists, elbows and ears. The authors claim that fluctuating asymmetry correlates between 0.20 and 0.25 with psychometric intelligence. Moreover, this is probably an underestimate — values of 0.30 to 0.35 might be expected in experiments using a better IQ test criterion and subjects from the full range of IQ, rather than the undergraduates from the University of New Mexico who were studied by Furlow and colleagues.

What form of stress might cause fluctuating asymmetry, and in what sort of sub-optimal development might it manifest itself? Furlow *et al.* cover all possibilities, including effects on any level of structure in the central nervous system. This is anything but satisfactory from a theoretical point of view, but quite in keeping with their report, which is essentially a data paper.

The authors propose that people are differentially susceptible to environmental

influences on their IQ, because of their genetic endowment. They believe that these susceptibilities are likely to have their effects early in life (for instance, prenatally). That is, they suppose that there is a real, common, causal link between bodily asymmetry and lowered IQ. Indeed, they are prepared to estimate that anything between 17 and 50 per cent of the variability in IQ is attributable to such causes. If the upper estimate proves to be accurate and replicable, then fluctuating asymmetry could account for almost all heritable sources of variability in IQ. But it is hard to believe that such a long-standing conundrum as the relative contributions of nature and nurture to IQ could be turned into no more than a neat paradox — nurture's influence depends on genetic predisposition — in one go.

The really neat thing is that Furlow *et al.* have used a method that is potentially within the reach of any psychology undergraduate — a method, in fact, that was almost within the reach of Francis Galton — requiring no more than callipers and the product-

moment correlation coefficient. They have a couple of tentative theoretical accounts as to why they should have got the results they report. These include structural developmental imperfections, leading to less efficient neural processing, and a differential use of energy budgets in people who have suffered varying degrees of developmental stress. But neither of these is more than an indicative hypothesis for the time being.

Expect many attempts at replication — low-life demonstrations that blacks/half-breeds/less-favoured races suffer more from fluctuating asymmetry than white, middle-class golden youth. Also expect, in general, a failure to recognize that Furlow and his colleagues may just have glimpsed a way of reconciling the long-standing antagonism between chiropodists and sky pilots. □

Steve Blinkhorn is at Psychometric Research and Development Limited, Brewmaster House, The Maltings, St Albans, Hertfordshire AL1 3HT, UK.

1. McClearn, G. E. *et al.* *Science* **276**, 1560–1563 (1997).
2. Furlow, F. B., Armijo-Prewitt, T., Gangestad, S. W. & Thornhill, R. *Proc. R. Soc. Lond. B* **264**, 823–830 (1997).

Galaxy evolution

The end of the beginnings

Roberto Abraham

An eerie feeling has overtaken recent gatherings of observational and theoretical cosmologists* — a sense of impending (if not quite actual) consensus. A broad outline of the processes by which galaxies form and evolve, in rough agreement with the predictions of theory, seems to have emerged from a number of recent galaxy surveys^{1–4}.

We classify galaxies by their distribution of stars, so galaxy evolution can be quantified by changes in star-formation rates. These can be measured by looking back to galaxies with increasing redshifts (thus further from us, and earlier in the Universe; Fig. 1), at different wavelengths, to see how the average brightness and spectrum change. In the past two years the accessible redshift range has expanded up to $z < 4$. This is remarkable progress: until recently, studies of 'normal' galaxy populations were restricted to redshifts $z < 1$. However, the discovery⁴ that the so-called 'Lyman dropout' galaxies (which appear anomalously faint through blue filters because a prominent ultraviolet spectral break has been redshifted into optical wavelengths) are at redshifts greater than two, has increased enormously the number of known high-redshift galaxies. Around 80–90% of the total age of the Universe is now being

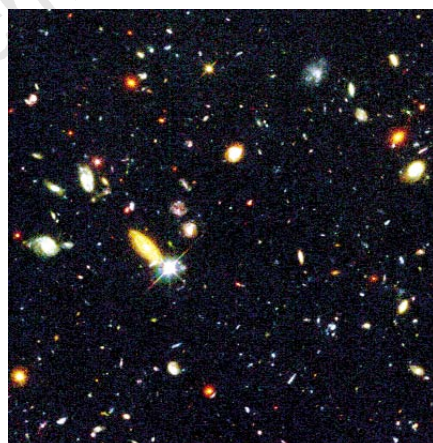


Figure 1 Back in time — the Hubble Deep Field, of which this is a view, has helped to revolutionize understanding of the Universe at high redshift.

probed by galaxy surveys. That is comparable to the fraction sampled by the relatively rare quasars and radio galaxies.

The current observational picture is summarized by the star-formation history curve (Fig. 2) (P. Madau, Space Telescope Sci. Inst.). By adding together the luminosity contributed by all the galaxies in a series of redshift intervals, this curve maps out a precipitous rise and fall in the production of stars and 'metals' (elements heavier than helium) in a given volume of the Universe as a function of redshift⁵.

The volume-averaged star-formation rate over such a large fraction of the total age

of the Universe is a powerful clue to the timescales over which galaxies form. The most remarkable aspect of Fig. 2 is the rapid decline at high redshift: this indicates that current surveys, by probing out to redshifts near 5, are now seeing almost all of the galactic history of the Universe. The epoch of peak star formation is around $z=1.5$.

The curve supports a hierarchical picture of galaxy formation. Hierarchical models "treat galaxy formation as a process, not an event" (S. White, Max Planck Inst. Astrophys.): the idea is that small, amorphous proto-galaxies form first, eventually settling into disk galaxies such as spirals, which can then also merge to form ellipticals. Rival theories suggest that giant galaxies form via the collapse of a single massive gas cloud at high redshift, with the nature of this initial collapse, rather than subsequent galaxy mergers, determining the form of the galaxy.

Hierarchical models predict star-formation histories that agree at least qualitatively with those in Fig. 2. There are a few theoretical embarrassments which refuse to disappear from the simulations — such as egregiously small galactic disks with low angular momenta (G. Efstathiou, Univ. Oxford; M. Steinmetz, Steward Obs.) — but even so there is considerable confidence in these models, and so much of the theoretical focus is now on the details of the different galaxy types^{6–8} that contribute light at different redshifts. For example, is the light at the highest redshifts coming from proto-elliptical galaxies (M. Giavalisco, Carnegie Obs.), or do ellipticals form continuously over a range of redshifts (G. Kauffmann, Max Planck Inst. Astrophys.)?

But there is one unresolved question about the overall star-formation rate. Some of it must be missing from Fig. 2 because of dust contamination. Dust scatters and absorbs light most strongly at short wavelengths, and, because of the expansion of the Universe and the corresponding redshift, the more distant the galaxy, the shorter was the original wavelength of the light we see. So it may be that distant, actively star-forming but very dusty 'starburst' galaxies contribute much of the total brightness of the Universe but are nonetheless missing from optical surveys. Sensitive infrared surveys will be able to detect such systems, if they exist.

The question is: how much light are we missing? Nearby starburst galaxies appear to have spectra that curve upwards from the ultraviolet to the infrared, and if dust-enshrouded starbursts in the distant Universe (at $z > 1$) have similar properties, they could contribute more than ten times as much energy as is accounted for in the star-formation history curve (G. Meurer and T. Heckman, Johns Hopkins). That agrees with a preliminary analysis of images obtained by the Infrared Space Observatory satellite,

*The Hubble Deep Field: Space Telescope Science Institute May Symposium, Johns Hopkins University, Baltimore, USA, 6–9 May 1997; and The Ultraviolet Universe at Low and High Redshift: Probing the Progress of Galaxy Formation, College Park, Maryland, USA, 2–4 May 1997.

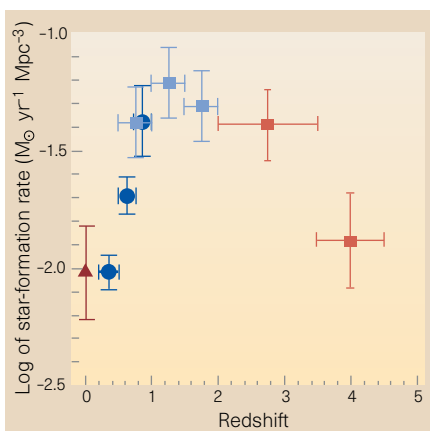


Figure 2 The star-formation history of the Universe. The formation rate is in units of solar masses per year per cubic megaparsec.

Individual points come from redshift surveys which, taken together, cover 80–90% of the total age of the Universe. (Red triangle from ref. 10; dark blue circles from refs 1 and 2; light blue squares from ref. 11; orange squares from ref. 5.)

from regions within the Hubble Deep Field (M. Rowan-Robinson, Imperial College). Furthermore, simulations of distant, ultra-luminous counterparts to local starburst galaxies show that some of them would resemble certain classes of peculiar galaxy detected in the Hubble Deep Field (W. Vacca, Univ. Hawaii).

But more modest dust extinction has been inferred from Lyman dropout systems (M. Pettini, Royal Greenwich Obs.; M. Dickinson, Johns Hopkins). Their spectra do not turn down strongly at short wavelengths, as would be expected from strong dust extinction. This still implies that the high-redshift tail of the star-formation history curve is too low, but only by a factor of 3–4.

There are two arguments in favour of the lesser degree of dust extinction. Increasing the curve by much more than a factor of 3–4 may result in a contradiction with existing data: the diffuse infrared background flux is close to that predicted from the present version of Madau's curve (L. Cowie, Univ. Hawaii). Also, to avoid producing more low-mass (therefore long-lived) stars than we see now, a factor of ten increase in the early star-formation rate would require that star formation at the time was biased in favour of producing only massive (hence short-lived) stars. But even this rather unpleasant assumption solves only part of the problem, because high-mass stars eject metals into the intergalactic medium, and the integrated abundance of metals from Madau's curve is roughly what we see in the local Universe. One possible way out is that the metals produced by high-redshift star formation are locked away unobserved in between galaxies⁹.

A final curiosity is that the star-formation history curve, at least in its present form, is

qualitatively similar to quasar evolution. Perhaps quasar evolution is more intimately linked to galaxy evolution than hitherto believed. If so, the star-formation history curve may help to address one of the great unsolved mysteries in astrophysics: where have all the quasars gone? □

Roberto Abraham is at the Royal Greenwich Observatory, Madingley Road, Cambridge CB3 0EZ, UK.

1. Lilly, S. J., Le Fevre, O., Hammer, F. & Crampton, D. *Astrophys. J.* **460**, L1–L4 (1996).
2. Ellis, R. S., Colless, M., Broadhurst, T., Heyl, J. & Glazebrook,

- K. *Mon. Not. R. Astron. Soc.* **280**, 235–251 (1996).
3. Cowie, L. L., Songaila, A., Hu, E. M. & Cohen, J. D. *Astron. J.* **112**, 834–864 (1996).
4. Steidel, C. C., Giallisco, M., Pettini, M., Dickinson, M. E. & Adelberger, K. *Astrophys. J.* **462**, L17–L22 (1996).
5. Madau, P. *et al. Mon. Not. R. Astron. Soc.* **283**, 1388–1404 (1996).
6. Giallisco, M., Steidel, C. C. & Macchetto, F. D. *Astrophys. J.* **470**, 189–194 (1996).
7. Baugh, C., Cole, S. & Frenk, C. S. **282**, L27–L32 (1996).
8. Abraham, R. G. *et al. Mon. Not. R. Astron. Soc.* **279**, L47–L52 (1996).
9. Mushotzky, R. F. & Loewenstein, M. *Astrophys. J.* (in the press).
10. Gallego, J. *et al. Astrophys. J.* **455**, L1–L4 (1995).
11. Connolly, A. J. *et al. Astrophys. J.* (in the press).

Potassium channels

Dendritic shock absorbers

Rafael Yuste

Mammalian neurons have dendritic trees — branched structures with complex geometries — which have mesmerized researchers for over a century. In 1891, Santiago Ramón y Cajal noticed that whereas sensory neurons point their dendrites towards the periphery, they always send their axons towards the brain. He suggested that a neuron receives information through its dendrites, and that it transmits this information through its axon. But exactly how the dendrites receive and integrate sensory inputs, and transfer them to the axon, is still a mystery.

On page 869 of this issue, Hoffman *et al.*¹ report that they have found potassium channels in dendrites from hippocampal pyramidal neurons. Why is this of general importance? As the authors show, potassium channels play a key role in dendritic excitability, and they may mediate sophisticated computations required for learning.

For most of this century, dendrites were considered to be passive cables. This view was shaken in the 1960s by reports of dendritic action potentials (electrical impulses) in cerebellar Purkinje cells² and pyramidal neurons³. The idea of active dendrites was slow to spread, perhaps because they were so difficult to work with (diameter < 5 μm). Recently, however, new optical, electrophysiological and immunocytochemical techniques have allowed a closer look at the function of dendrites. And these studies have revealed sodium and calcium channels in the dendrites from many different types of neurons^{4,5}.

Still missing from this picture were the potassium channels. As Hille⁶ wrote, “potassium channels are like the stops of an organ”, because their rich, functional diversity creates many different ways of dampening electrical excitation and setting the timbre of neurons. For instance, potassium channels

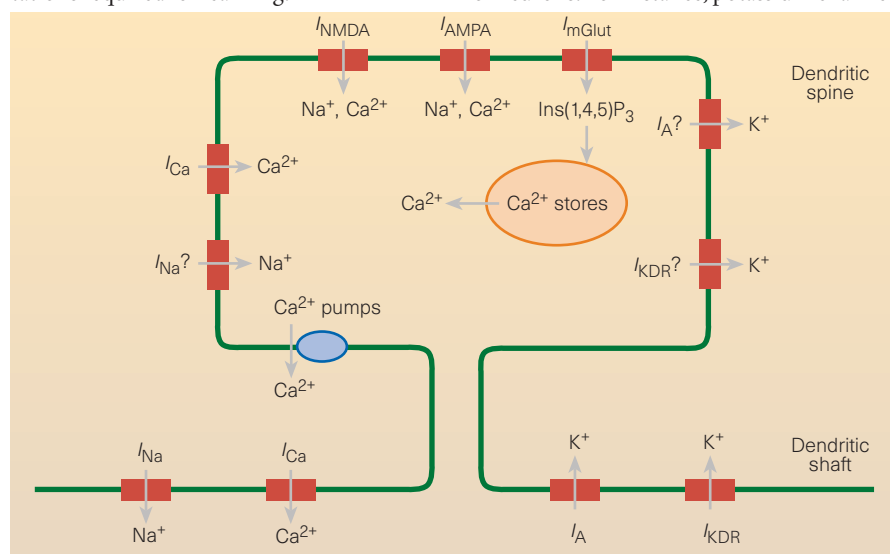


Figure 1 Known and postulated channels and receptors in dendritic spines and dendritic shafts of CA1 neurons in the hippocampus. Conductances: I_A , A-type potassium channels; I_{KDR} , delayed-rectifier potassium channels; I_{Na} , sodium channels; I_{Ca} , calcium channels. Glutamate receptors: I_{NMDA} , N-methyl-D-aspartate receptor; I_{AMPA} , aminohydroxymethylisoxalate propionic acid; I_{mGluT} , metabotropic glutamate receptor. Ins(1,4,5)P₃, inositol-1,4,5-trisphosphate.

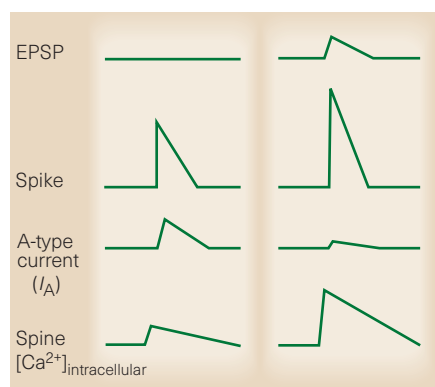


Figure 2 Hoffman *et al.*¹ have found that potassium channels exist on dendrites of hippocampal pyramidal (CA1) neurons. They propose that the temporal coincidence of excitatory postsynaptic potentials (EPSPs) and back-propagating action potentials (spikes) results in a larger spike, because these dendritic A-type potassium channels are inactivated. The larger spike will lead to a greater influx of calcium into the dendritic spines, and this may underlie long-term potentiation — a putative cellular mechanism of learning.

maintain the resting potential, and they regulate the hyperpolarization of an action potential. So, in dendrites, potassium channels could act as 'shock absorbers', dampening the depolarization that is produced by excitatory inputs — also known as excitatory postsynaptic potentials (EPSPs) — and the effect of the sodium channels (which produce action potentials) and calcium channels. Although dendritic potassium channels have been found in turtle Purkinje cells⁷, until now the existence and possible function of potassium channels in mammalian dendrites had not been directly shown.

Hoffman *et al.*¹ have done a detailed study of dendritic potassium channels in apical dendrites of CA1 pyramidal neurons from the hippocampus, and they find a conductance with physiological and pharmacological profiles corresponding to the so-called A-type potassium channels. The authors also describe lower levels of another conductance that is analogous to the delayed-rectifier potassium channel. Finding a strong A-type current in apical dendrites is consistent with previous studies using intradendritic recordings⁸ and immunohistochemistry⁹, increasing the number of voltage-sensitive channels that are known to be present in CA1 dendrites. So, not only are dendrites active, but they may have more different types of channels than axons (Fig. 1).

But what are all of these active conductances doing? In many types of neuron, action potentials from the cell body are known to 'back-propagate' through the dendritic tree¹⁰ (although the extent of this phenomenon *in vivo* is still unclear¹¹). This back-propagating action potential (or spike) is mediated by dendritic sodium channels.

It reaches the fine dendritic branches and spines, opening calcium channels along its way and causing large influxes of calcium into dendritic shafts and spines^{12,13}. So, dendritic spines, which receive the EPSPs (that is, the input into the cell), can also 'hear' the spike (that is, the output of the cell), meaning that they are ideally poised to detect and compute temporal coincidences of neuronal inputs and outputs.

In neocortical pyramidal neurons, the temporal coincidence of EPSPs and back-propagating spikes produces long-term synaptic potentiation (LTP) — a cellular mechanism that may underlie learning^{14,15}. But if the spikes reach the dendritic spines before the EPSPs occur, there is a long-lasting decrease in synaptic efficacy; that is, cellular 'forgetting'¹⁵. So, back-propagating spikes, and their exact timing with respect to the EPSPs, are required for this neuronal plasticity.

In CA1 neurons, back-propagating spikes also seem to be necessary for LTP, because LTP does not occur if sodium-channel-mediated spikes are blocked^{16,17}. The detection of coincident EPSPs and back-propagating spikes in LTP is thought to be implemented by an influx of calcium through the NMDA (*N*-methyl-D-aspartate) type of glutamate receptor (Fig. 1) when glutamate binds to it and the postsynaptic cell is depolarized. This calcium influx could then orchestrate long-lasting changes in synaptic efficacy. Indeed, LTP does not occur if NMDA receptors are blocked or if postsynaptic calcium is chelated¹⁴. In agreement with the idea that the dendritic spines detect coincident EPSPs and back-propagating spikes, when the two phenomena occur simultaneously, higher, supralinear calcium accumulates in the spines¹³.

Dendritic A-type currents now provide a new mechanism for this detection. Using computer simulations, Hoffman *et al.*¹ find that dendritic A-type channels can be inactivated by the depolarization that is produced during EPSPs. This inactivation will produce a larger, full-blown, back-propagating spike, because the A-type channels will no longer dampen it (Fig. 2). Indeed, a larger back-propagating spike is measured in apical dendrites when an EPSP occurs simultaneously¹⁷. This larger spike could open more calcium channels in the dendritic spine, increasing the influx of calcium during coincident EPSPs and spikes. Moreover, the larger back-propagating spike would unblock the NMDA receptors more efficiently, increasing their permeability to calcium. So the temporal coincidence of the input and output of the cell would be detected by the effect of the A-type channels on the spike, and translated into a greater accumulation of calcium. But more experiments are needed to establish this interesting hypothesis.

LTP is, in fact, only one of the many functions that occur in dendrites. Dendritic potassium channels could influence other aspects of dendritic integration; for example, active dendritic conductances could shape the EPSPs. Indeed, Hoffman *et al.* show that dendritic sodium channels boost EPSPs, but that this boost is normally prevented by A-type potassium currents. Finally, potassium channels could also control dendritic spiking, because blocking A-type currents leads to the generation of dendritic action potentials triggered by EPSPs or an injection of current⁸. So the spike initiation zone, which is generally assumed to be in the axon and/or cell body, may shift to the dendrite if potassium channels are inactivated under particular physiological conditions. Dendritic potassium channels may also help us to understand why spikes are not generated in the dendrites — where EPSPs originate — in spite of there being a similar density of sodium channels in dendrites and the cell body.

Potassium channels have many regulatory mechanisms. All of the putative functions of the dendritic potassium channels could be specifically regulated in different parts of the dendritic tree. For example, trains of EPSPs or IPSPs (inhibitory post-synaptic potentials), second messengers, modulatory neurotransmitters and changes in extracellular ions or in redox states of the cell could, in principle, modify the activation or inactivation kinetics of potassium channels and alter the delicate balance of active conductances in the dendrites. Now, more than ever, a detailed knowledge of the nuts and bolts of dendritic processing seems essential for a proper understanding of neuronal plasticity and function. □

Rafael Yuste is in the Department of Biological Sciences, Columbia University, 1212 Amsterdam Avenue, New York, New York 10027, USA.

- Hoffman, D. A., Magee, J. C., Colbert, C. M. & Johnston, D. *Nature* **387**, 869–875 (1997).
- Llinás, R., Nicholson, C., Freeman, J. A. & Hillman, D. E. *Science* **160**, 1132–1135 (1968).
- Spencer, W. A. & Kandel, E. R. *J. Neurophysiol.* **24**, 272–285 (1961).
- Johnston, D., Magee, J. C., Colbert, C. M. & Christie, B. R. *Annu. Rev. Neurosci.* **19**, 165–186 (1996).
- Yuste, R. & Tank, D. W. *Neuron* **16**, 701–716 (1996).
- Hille, B. *Ionic Channels in Excitable Membranes* (Sinauer, Sunderland, MA, 1992).
- Hounsgaard, J. & Mitgaard, J. *J. Physiol., Lond.* **402**, 731–749 (1988).
- Andreasen, M. & Lambert, J. D. C. *J. Physiol., Lond.* **483**, 421–441 (1995).
- Sheng, M., Tsaur, M., Jan, Y. N. & Jan, L. Y. *Neuron* **9**, 271–284 (1991).
- Stuart, G. J. & Sakmann, B. *Nature* **367**, 69–72 (1994).
- Svoboda, K., Denk, W., Kleinfeld, D. & Tank, D. W. *Nature* **385**, 161–165 (1997).
- Spruston, N., Schiller, Y., Stuart, G. & Sakmann, B. *Science* **286**, 297–300 (1995).
- Yuste, R. & Denk, W. *Nature* **375**, 682–684 (1995).
- Baudry, M. & Davis, J. L. *Long-Term Potentiation* Vols 1–3 (MIT Press, Cambridge, MA, 1991, 1994, 1997).
- Markram, H., Lübke, J., Frotscher, M. & Sakmann, B. *Science* **275**, 213–215 (1997).
- Scharfman, H. E. & Sarvey, J. M. *Brain Res.* **331**, 267–274 (1985).
- Magee, J. C. & Johnston, D. *Science* **275**, 209–212 (1997).

Nonlinear optics

Solitons light the way

Allan Boardman

Moti Segev and his team from Princeton University were the first into the field of self-trapped (soliton), coherent light beams in photorefractive materials. Not content with that, on page 880 of this issue Mitchell and Segev now announce the trapping of incoherent light beams (*Nature* 387, 880–883; 1997). This result lends great flexibility to their pioneering work, and points the way towards the development of devices based on the principle. An appreciation of this advance demands knowing what a soliton is, and why photorefractive materials are so useful, which in turn requires stepping back in time.

One of the most significant observations in science was made 153 years ago by a naval architect, called John Scott Russell. This was his sighting of a very peculiar wave on the Edinburgh to Glasgow canal. For some reason, a boat he was observing stopped suddenly and caused a hump of water to

accumulate around its front. This hump then took off down the canal. Russell, a man blessed with acute observational powers, soon noticed that this heap of water was nicely rounded and, rather strangely, propagated without change of shape. The hump had no disturbance in front of it and none behind it. It was, in fact, alone (that is, solitary) and did not conform to common sense by dying away, as wave disturbances caused by striking water with a stone, or in this case a boat, are expected to do.

It took two Dutchmen, Korteweg and de Vries, in 1895, to realize that the key feature was that the hump of water had an unusually large amplitude so that a nonlinear restoring force needed to be invoked (although, not surprisingly, Rayleigh had come close to this theory in 1876). Normally, a hump of water like this is thought of as a packet of waves, all travelling with different speeds. This way of looking at things comes from Fourier, but it

is a linear viewpoint and the end result is the destruction of the hump. Large amplitudes mean lots of power, so the waves in the packet are now forced to interact with one another. The forces trying to restore equilibrium are no longer just proportional to the height of the wave and the speed now depends on displacement: this is a nonlinear system. In a balanced situation, the power of the wave acts against dispersion. The wave remains intact and a soliton is born!

As is sometimes the way of science, this nonlinear work was then ignored as being of marginal interest. Indeed, the natural philosophers of the time remained blissfully unaware that our world is really a nonlinear one. Russell was convinced of the importance of his wave, and its curiosity, right up to his death in 1882. But it took much later research on waves in crystal lattices, by Zabusky and Kruskal in 1965, to make plain the elegant generic nature of Russell's observation to a wider public, and that nonlinearity is important to the propagation of such waves; Zabusky and Kruskal realized that the form of the equations concerned was exactly like that of Korteweg and de Vries's.

The really striking feature of the 1965

Palaeoclimatology

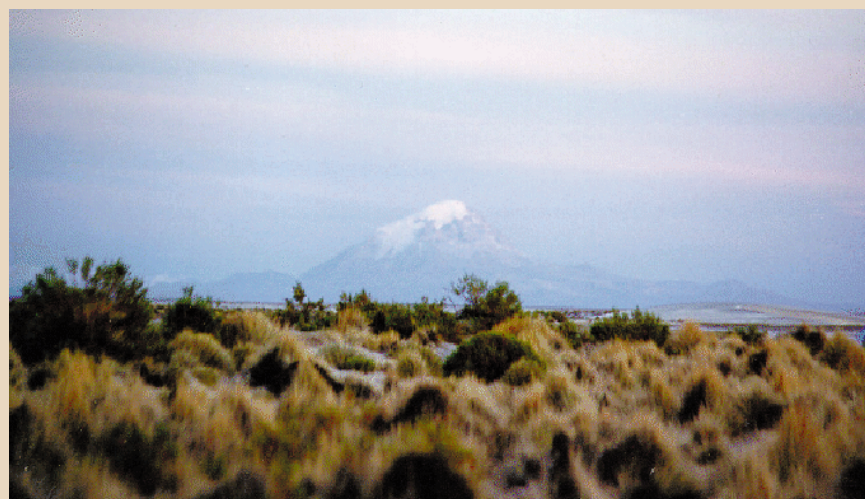
Hot air, big chill

Now here's a teaser. You want to get three 115-m-long ice cores down from a glacier on a tropical volcano without them melting. What's more, you want to keep the ice colder than -15°C , and the only way down to the more typically tropical high-plateaus below is by foot and beast of burden.

One solution is to carry a hot-air balloon up the mountain in 14 backpacks, assemble it atop the peak, and fly sections of the core to the nearest freezer. Which is exactly what Lonnie Thompson and colleagues, from the Byrd Polar Research Center at Ohio State University, are about to do on Sajama (pictured here), a 6,548-m-high stratovolcano at 18°S , 68°W on the Bolivian plateau.

The reason for carrying out this exercise is that glaciers contain some of the most reliable indicators of past climate. The chemical and isotopic composition of the ice, as well as trapped pollen and dust particles, have told us much of what we know about past atmospheric composition and climate change over the past hundred thousand years or so. Sajama should provide the highest-resolution, good-quality ice record that the tropics have yet yielded, possibly stretching back some 20,000 years.

Even more appealingly, tiny air bubbles trapped in the ice can provide samples of gases from past atmospheres, a vein of



information that has been successfully mined at high latitudes. A trapped-gas record from low latitudes should reveal much more about the timing and, ultimately, the mechanisms of past climate change, but if the ice becomes warmer than about -15°C a slight melting might cause the air bubbles to leak their story away. Todd Sowers (Penn State University) hopes to retrieve the gas record intact from the Sajama cores, and measure the methane concentrations and the isotopic composition of the oxygen and nitrogen.

Maintaining the big chill isn't too tricky in polar regions, but meeting the challenge

in the tropics first means getting high enough to find old, cold ice, and then keeping it cold — which is where of course the hot-air balloon, the *Soaring Penguin*, comes in. The ice core will be drilled using solar energy and sectioned into foam-insulated boxes. This precious cargo will then be strapped to the balloon's basket and flown to the base-camp freezers 2,000 m below. All that remains is to truck the freezers down the lower slopes of the mountain and transport them back to the laboratories in the United States. As those who have done fieldwork know, that may prove to be the hardest job of all.

Philip Newton

work, however, was in demonstrating that solitary waves retain their shape (this time in a crystal), even after colliding with each other. The authors invented the name soliton for these waves, just to emphasize that although a soliton is a solitary wave it retains its identity, even after a collision. The change of 'ary' to 'on' is not surprising, incidentally — physicists call things 'ons' at almost every opportunity (take electron, meson and photon, for instance) because the Greek word 'on' means solitary, and in each instance the suffix signals the principle that the entity concerned retains its characteristics.

We are well aware that a beam of light, even if it is emerging from the now familiar laser pointer used by lecturers, will spread out and change shape as it propagates. In air, this is all that will happen. But what about light propagation in other media, especially photorefractive materials — which, importantly, already have a mature fabrication technology to back them up? In this case light always heads for the region of highest refractive index (think of how a lens works), and a photorefractive material such as lithium niobate is special because the light grabs hold of the electrons and forces the index up, precisely and only where the beam is travelling. This is just what is needed to stop light rays from inside the beam from bending outwards and causing a change in beam shape. When beam spreading is stopped by the photorefractive material, a spatial soliton is formed and the beam can now be used as a stable optical guide, just like the familiar optical fibre.

The spatial soliton is, however, better than the fixed, 'hard' waveguide such a fibre constitutes. The reason is that it is a 'soft' guide that is easily tunable by changing the beam size. Furthermore, these soft guides can be set up at a millionth of a watt of power in most photorefractive materials. Having done this, the incredible thing is that tasks can be performed inside the guide. For example, another light beam can be sent down it at a million times this power, provided its colour forbids the grabbing of more electrons to change the guide, size or direction. Many more applications can be imagined for these novel waveguides, such as using them for optical wiring, to steer one beam with another in addressable arrays, and to act as switching devices and as logic units for a whole range of information processing and computing.

The reason why the new work of Mitchell and Segev is really compelling is because they have come up with a new soliton just when we have all become used to using coherent (laser) beams in which the phase at one point determines the phase at all points. Spatial solitons — self-trapped beams — which have randomly varying phase yet maintain a smooth, well-defined, constant profile in photorefractive materials, are a

truly significant outcome of modern soliton and materials science because of the ease with which they can be manipulated. Although there are still worries about how fast the photorefractive materials can react to light, or changes of beam direction, the discovery should herald an era of new, all-optical processing devices that are cheap and easy to implement. Russell's observation did

not change physics or technology in his own century nor most of our own. But with the real possibility of optical soliton communication systems and all-optical processors now on the horizon, it may well transform them next century. □

Allan Boardman is in the Joule Laboratory, Department of Physics, University of Salford, Salford M5 4WT, UK.

Theoretical biology

A robust view of biochemical pathways

Leland Hartwell

Stanislaw Leibler, a physicist from Princeton University, and his student, Naama Barkai, have taken a fresh look at the design of a biochemical pathway. They have formulated a new concept, robustness, which has fundamental implications for the tolerance of genetic polymorphisms in outbred populations (page 913 of this issue: *Nature* 387, 913–917; 1997). Biochemical pathways are composed of groups of proteins that cooperate to transmit signals (as in the induction of gene expression by an extracellular hormone) or to function as machines (as in DNA replication).

Barkai and Leibler consider the chemotaxis pathway of bacteria, a set of 12 proteins that receives signals from the extracellular environment and transmits these signals to the molecular machinery that controls the motion of the bacterial flagella. This pathway allows bacterial cells to swim up gradients of attractants (such as food) or down gradients of repellents (such as poisons). Barkai and Leibler pose the following questions. Does the pathway work the way it does because evolution has fine-tuned the kinetic parameters and amounts of proteins to achieve the desired behaviour? Or is there something more fundamental at work in the construction of the pathway — something that is insensitive to fine-tuning, which they term 'robustness'.

To tackle these questions, they develop a computer simulation of the chemotaxis pathway and substitute altered kinetic constants and protein concentrations to see what effects these changes will have on the behaviour of the system. After thousands of simulated experiments they come to the conclusion that one aspect of the chemotactic response, the ability of the system to adapt to its basal state of response after perturbation, is robust. That is, gross changes in many parameters do not alter the ability of the system to adapt precisely (in this case for the frequency with which the bacterial flagella switch their direction of rotation to return to the frequency characteristic of no signal). More than 80 per cent of simulations in



Figure 1 *Escherichia coli* and its flagella in false-colour glory. Barkai and Leibler's concept of 'robustness' stems from their simulations of the bacterial chemotaxis signalling system ($\times 8,000$).

which every parameter had been changed randomly by a factor of two, or simulations in which any one parameter had been changed by several orders of magnitude, displayed nearly perfect adaptation. The robustness of adaptation is accounted for as a design principle of the pathway — one output of the signalling system (protein methylation) is a negative feedback circuit that reduces signal from the receptors.

Now, the types of changes that the chemotactic pathway is insensitive to are the types of changes that would occur naturally in an outbred population — natural polymorphisms that affect kinetic parameters or amounts of proteins. As a geneticist working with the yeast *Saccharomyces cerevisiae*, I am familiar with the curse of genetic polymorphisms. A particular mutation can be lethal in one background and have no apparent effects in another. The profound effects of genetic background lead one to wonder how outbred populations are able to tolerate all

the combinations of polymorphisms that exist in the population.

The concept of robustness suggests an answer. Biochemical pathways are constructed in such a way that their output is relatively insensitive to the types of variations that occur in natural polymorphisms. One might say that this result is inevitable, because any polymorphisms that frequently resulted in combinations that were lethal or caused infertility would be eliminated. However, Barkai and Leibler did not limit themselves to naturally occurring polymorphisms. If their simulations apply to the real world, their results imply that at least this particular pathway is insensitive to almost all polymorphisms.

How can we begin to think about the implications of genetic diversity on the biology of outbred organisms? Simple concepts that come to mind are those of complex traits and genetic modifiers, ideas that ultimately centre on one or a few genes. Rather than focusing attention on individual proteins, Barkai and Leibler suggest focusing attention on the collection of proteins that constitute a biochemical pathway. To generalize their concept, one could imagine that all pathways are constructed to be insensitive to fine-tuning. Although this hypothesis is certainly not justified by consideration of only one pathway, and that by simulation only, the idea leads to many interesting questions that are ultimately subject to experiment.

Obviously, the first question is what the consequences are if diversity is generated in a real cell. A second is the issue of what constitutes a pathway. The chemotaxis pathway considered by Barkai and Leibler is composed of 12 proteins sitting in a cytoplasmic sea of several thousand other proteins. Is the set that we consider a pathway the same as the set that evolution has considered a pathway? How could we tell? A result of Barkai and Leibler's suggests a test. They found that the ability to adapt to basal response was robust but that the time required to adapt was not. The set of proteins that constitute one pathway and affect one behaviour might be tested to see if they are part of or distinct from the set that constitute another pathway (and affect a second behaviour), in the following way. Introduce variation in one set of proteins that perturbs a non-robust parameter of behaviour 1; then determine whether this change affects a non-robust parameter of behaviour 2.

A further question is what the robust parameters of other pathways might be. An answer would identify the most important output of a pathway, the characteristic that was selected by evolution to be robust. Finally, do apparently redundant pathways exist because one has been selected to be robust for one output whereas another has been selected to be robust for a different output, thereby allowing the cell to be robust to both

outputs? It might be fruitful to consider genetic disease as due to those rare alleles that map outside the zone of robust variation.

Readers should turn to the paper itself on page 913 to find out in more detail about the authors' ideas. I am sure they will think of

many more questions that can now be formulated as a result of this powerful concept of robustness. □

Leland Hartwell is in the Fred Hutchinson Cancer Research Center, 1124 Columbia Street C2-023, Seattle, Washington 98104, USA.

Biological oceanography

Sulphur, climate and the microbial maze

Gillian Malin

It might seem strange that marine algae and climate are linked, but there is a connection and it occurs through the production of a compound which breaks down to release volatile dimethyl sulphide — DMS, $(\text{CH}_3)_2\text{S}$ — and acrylic acid. Formation of the volatile trace gas has implications that go far beyond the level of an individual cell or microbial population, because DMS is a key compound in the global sulphur cycle and its oxidation products influence atmospheric acidity, as well as cloud formation and the Earth's temperature¹. The compound in question is dimethylsulphoniopropionate (DMSP), which algae synthesize to help maintain their

osmotic balance with sea water.

Papers on pages 891 and 894 of this issue advance our understanding of how DMS is produced in the marine environment. Gage *et al.*² reveal the biochemical pathway for DMSP synthesis in a diverse group of marine algae. Wolfe and co-workers³ provide fresh insight into the involvement of grazing microzooplankton in releasing DMSP from marine phytoplankton cells.

The main sources of DMSP in the natural environment are marine phytoplankton blooms, where high levels of primary production and biomass occur on a timescale of weeks (Fig. 1), and phytoplankton assem-

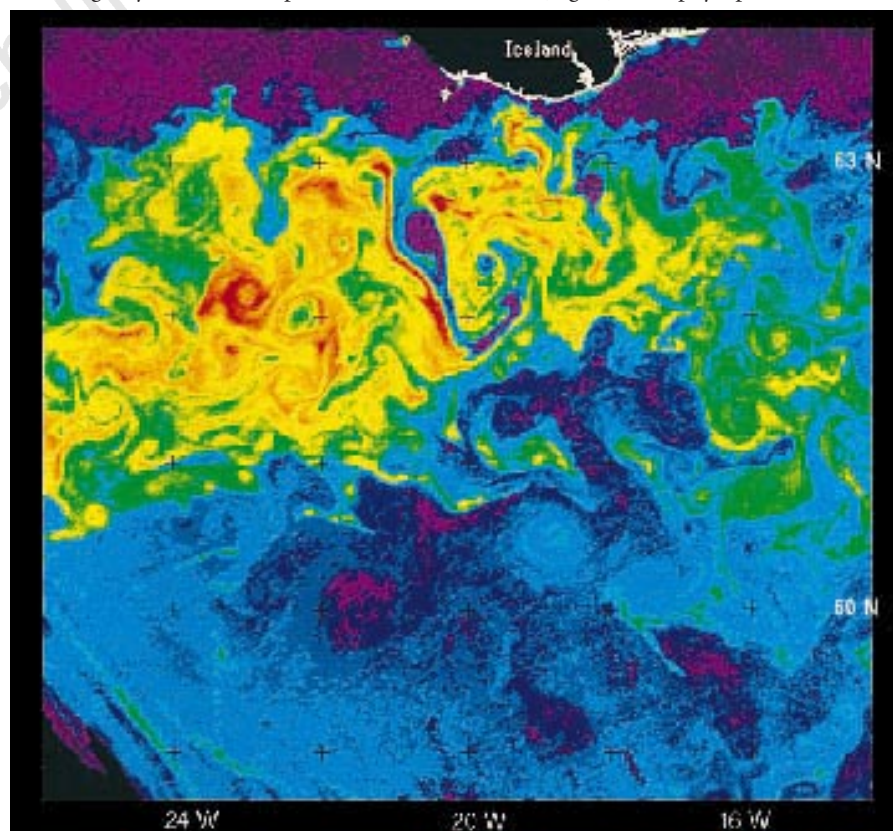


Figure 1 False-colour satellite image depicting the reflectance due to a large-scale bloom of *Emiliana huxleyi* in the northeast Atlantic, south of Iceland, in June 1991. Blooms such as this produce large amounts of dimethyl sulphide (DMS), the oxidation products of which influence atmospheric chemistry and climate. The rainbow colour scale represents high (red–orange) to low (blue–violet) levels of reflectivity of the upper layer of the sea. The reflectivity, which intensifies towards the end of a bloom, is due to the cells shedding coccoliths into the water column as they grow. (Image processed by S. Groom, Remote Sensing Data Analysis Service, Plymouth Marine Laboratory.)

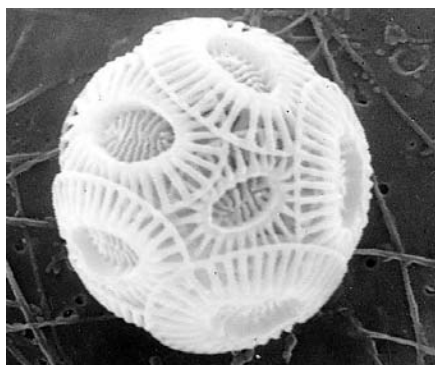


Figure 2 Scanning electron micrograph of a cell of the marine phytoplankton species *Emiliana huxleyi*, approximately 5 μm in diameter. The outside of the cell is covered with ornate interlocking plates of calcium carbonate known as coccoliths. This organism contains high concentrations of dimethylsulphoniopropionate (DMSP), a precursor of DMS. The work discussed here^{2,3} shows the route by which DMSP is synthesized in marine algae, and how grazing results in its conversion to DMS and acrylic acid. (Micrograph courtesy of D. Harbour, SAHFOS, Plymouth.)

blages in the vast areas of the open oceans where steady-state populations exist for longer periods. Every millilitre of sea water contains a large and diverse microbial population of photosynthetic cells, zooplankton, bacteria and viruses. The processes by which DMSP is released from phytoplankton cells and DMS is produced are best depicted by a network of production, transformation and consumption which involves the entire marine microbial food web⁴. To date several key processes have been identified, although turnover rates and what controls them are not well understood. What is clear is that less than 10 per cent of the DMS in surface sea water ever enters the atmosphere. The rest is turned over by bacteria or oxidized by sunlight to form non-volatile products.

Grazing is a major route for the release of DMSP to sea water. In many regions unicellular microzooplankton are the principal herbivores, and they have a particularly important impact on small marine phytoplankton cells such as *Emiliana huxleyi* — a species that contains large amounts of DMSP (Fig. 2). These protozoa are highly selective in what they eat, and use chemical cues to select and reject prey⁵. Some DMSP-containing microalgae have an enzyme known as DMSP lyase, which does not come into contact with intracellular DMSP during normal growth. Wolfe *et al.*³ propose that grazing mixes DMSP with the enzyme, thereby liberating DMS and acrylic acid, and then go further to suggest that DMSP may act as a defence precursor. Acrylic acid has signalling properties for a range of organisms, including bacteria, protozoa and fish, and the herbivorous microzooplankton used in the study actively retreated from gradients

of the compound. In the sea this response might shift grazing pressure away from cells that contain high levels of DMSP. Wolfe and colleagues' report is the first of grazing-activated defence for a unicellular organism.

In their paper, Gage and co-workers² describe a series of elegant biochemical experiments through which they have identified the intermediates in DMSP biosynthesis, including the novel compound dimethylsulphoniohydroxybutyrate (DMSHB). They show that the pathway in a common green seaweed is different from that found in higher plants, which suggests that these pathways evolved independently. However, macroalgae, which are typically found in the intertidal and subtidal zones of rocky shores, are thought to have only a small effect on the global sulphur cycle when compared with marine phytoplankton. With this in mind, Gage *et al.* went on to identify DMSHB in three phytoplankton species from different algal classes. DMSHB was readily catabolized when added to algal cultures, and the authors make the intriguing suggestion that it could be an additional precursor of DMS *in vivo*.

Marine organisms usually have a combination of organic osmolytes. DMSP is similar in both structure and function to the nitrogen-containing osmoregulatory compound glycine betaine (GBT). It has been proposed that, under conditions of nitrogen limitation⁶, a common occurrence in the marine environment, organisms might preferentially change to making the sulphur-containing DMSP molecule. There is evidence from laboratory studies and a field study that such a

GBT–DMSP switch exists⁷, and Gage and colleagues now provide further support for the idea at the biochemical level. DMSP synthesis is initiated by a transamination reaction, and this pathway would be favoured in nitrogen-depleted cells, hence effectively favouring a switch to DMSP.

And what of the climate link? Global climate change is a big issue⁸, and there is general awareness of how human activities are changing the concentration and distribution of greenhouse gases. It is less well known that atmospheric aerosols, which are submicrometre particles or droplets, have a cooling effect on global temperature. The link with algae is that some of the marine DMS escapes into the atmosphere where it oxidizes rapidly to form aerosol particles; these absorb and reflect solar energy directly back into space, and scatter radiation indirectly by acting as cloud-condensation nuclei.

Natural aerosol levels have been substantially affected by anthropogenic sulphur, especially in the Northern Hemisphere. Nevertheless, DMS oxidation products remain the major source of aerosol particles over remote marine areas (that is, over most of the globe). Anything that disturbs the balance of the microbial ecology of DMS production, such as changes in wind speed that alter the upwelling of nutrients, or increases in sea-water temperature that change the spatial distribution of grazers relative to their phytoplankton prey, could affect how much of this sulphur compound is emitted to the air, and therefore affect climate.

In 1987 Charlson and colleagues¹ sparked a flurry of research interest by proposing that

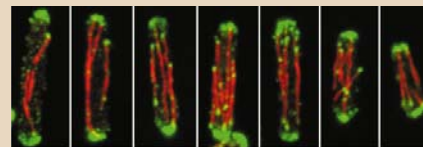
Cell biology

Explorers deliver tea to the pole

A few weeks ago, a report in these pages claimed that drinking green tea can prevent cancer (see *Nature* 387, 561; 1997). Now, in the 13 June issue of *Cell*, Juan Mata and Paul Nurse report (*Cell* 89, 939–949; 1997) that another type of tea could help you to grow properly — if you're a fission yeast, that is.

Immediately after cell division, the fission yeast *Schizosaccharomyces pombe* begins to grow. Initially, it extends only at the old tip that existed in the mother cell. But at a certain point in the cell cycle, it switches to bipolar growth, and its ends elongate directly away from one another. How does it do this?

To find out, Mata and Nurse studied yeast mutants in which the tips were not precisely opposed. Some of these mutants were bent, others were T-shaped, but all had defects in one of two *tea* (tip elongation aberrant) genes. Having cloned *tea1*, the authors found that the *tea1*



protein is tightly localized to both ends of a cell. Not only that but, using immunofluorescence staining, they saw that *tea1* (green) colocalizes with the ends of tubulin (red) microtubules.

Mata and Nurse believe this to mean that *tea1* acts as an end marker, directing the yeast's growth machinery to the cell poles. *Tea1* is delivered to the poles by the microtubular cytoskeleton, the organization of which may be affected as a result. This produces a dynamic system in which the space of the cell is explored, then the two growing tips are identified and maintained by *tea1* — a kind of tea for two for yeast.

Alison Mitchell

the connection between algae and clouds represents not only a climate linkage, but also a climate-regulating mechanism. Since then, as recent papers attest^{9,10}, there has been a great deal of progress, and evidence in support of the climatic significance of DMS continues to grow. For instance, clear relationships have been shown to exist between high levels of DMS in the atmosphere, increased particle concentrations and enhanced cloud albedo. In all, it is clear that the myriad interactions between the components of the microbial food web and their physical and chemical environment fundamentally influence how much DMS is available for exchange with the atmosphere. But direct proof of the phytoplankton–DMS–climate feedback idea, whereby climate would affect the quantity of DMS emitted to the air, remains elusive.

The papers by Gage *et al.*² and Wolfe *et al.*³ are examples of how, by increasing our understanding of the marine microbial maze,

we can inch closer to the ultimate goal of being able to predict how changes in environmental parameters might impinge on natural inputs of sulphur to the atmosphere. □

Gillian Malin is in the School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

e-mail: g.malin@uea.ac.uk

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Gage, D. A. *et al.* *Nature* **387**, 891–894 (1997).
3. Wolfe, G. V., Steinke, M. & Kirst, G. O. *Nature* **387**, 894–897 (1997).
4. Malin, G., Liss, P. S. & Turner, S. M. in *The Haptophyte Algae* (eds Green, J. C. & Leadbeater, B. S. C.) 303–320 (Clarendon, Oxford, 1994).
5. Verity, P. G. J. *Protozool.* **38**, 69–76 (1991).
6. Andreae, M. O. in *The Role of Air–Sea Exchange in Geochemical Cycling* (ed. Buat-Menard, P.) 331–362 (Reidel, Dordrecht, 1986).
7. Liss, P. S., Hutton, A. D., Malin, G., Nightingale, P. D. & Turner, S. M. *Phil. Trans. R. Soc. Lond. B* **352**, 159–169 (1997).
8. Houghton, J. T. *et al.* (eds) *Climate Change 1995: The Science of Climate Change* (Cambridge Univ. Press, 1996).
9. Kiene, R. P., Visscher, P. T., Keller, M. D. & Kirst, G. O. (eds) *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (Plenum, New York, 1996).
10. Cox, R. A. *et al.* (eds) *Phil. Trans. R. Soc. Lond. B* **352**, 139–254 (1997).

Gamma-ray astronomy

Bursting out all over

Joshua S. Bloom and Malvin Ruderman

Enigmatic bursts of γ -rays from cosmic sources are detected about once a day from every direction on the sky. During their brief duration, typically between 10^{-2} and 10^3 seconds¹, these bursts often have the brightest γ -ray intensities ever seen. Now, after almost three decades of observation and speculation, the distance scale to γ -ray burst (GRB) sources seems at last to have been settled. Evidence has been growing at a remarkable rate in the past few months — the most striking pieces appearing on pages 876 and 878 of this issue^{2,3} — to support the proposal that most, and perhaps all, of the bursts come from explosions involving neutron stars or black holes near the edge of our observable Universe. If this is indeed the case, GRB sources are among the most powerful events in the Universe. This was the view of most of the participants in a GRB workshop last month*.

In 1991, the Burst and Transient Source Experiment (BATSE), on board the Compton Gamma-Ray Observatory, began to accumulate what has now grown to almost 2,000 burst observations. These show us to be at the centre of a nearly homogeneous distribution of bursting sources, and the paucity of faint bursts implies that we are seeing almost to the edge of the source population⁴. These two features are natural consequences of any uniform source distribution in the simplest cosmological models for our expanding Universe⁵. Other suggestions, for

example that GRBs are from relatively small, recurrent explosions on the surface of neutron stars now in the halo of our own Galaxy⁶, needed some *ad hoc* tuning of parameters such as delays in the onset of explosions. This made such models less appealing, even though they more easily explained certain controversial features of some bursts (such as cyclotron resonance spectral features, periodicities of several seconds, and possible burst source repetitions).

A 'smoking gun' was needed, to identify some GRB source with a familiar type of object which could determine the distance-scale. The problem was that γ -ray telescopes such as BATSE have an angular resolution of several degrees, and so cannot pin down a burst location finely enough for other kinds of telescopes to identify a unique counterpart at other wavelengths. The Interplanetary Network of satellites, started in the 1970s, could triangulate the angular position of a burst from arrival-time differences at different satellites, but gave only a handful of arc-minute-sized error boxes, generally months after a burst — far too long an interval to allow the identification of other transient emissions associated with it.

Since its launch in 1996, the Italian–Dutch satellite BeppoSAX has improved on both position and time constraints enormously: it can locate X-ray emission from a GRB to a few arc minutes, and then tell optical and radio telescopes where to look within a few hours. So far, it has identified three fading X-ray sources as probable GRB afterglows. The 28 February 1997 burst⁷ was



100 YEARS AGO

'Styles of the calendar' — At the approach of the end of a century, this subject naturally comes to the front again; but it has lately been somewhat unexpectedly raised to special prominence by the suggested probability of one at least of the Oriental countries of Europe adopting the usage which, on the initiative of Rome in 1582, all the western nations gradually accepted, England (we say advisedly England *not* Britain, because Scotland adopted it before the union even of the crowns) being the last in 1752. America having been colonised by the western Europeans, and the United States having still been British colonies at the date last mentioned, the Gregorian style is universal in that continent. But eastern Europe, including Russia and all the nationalities of the Balkan peninsula, still adheres to the old Julian style; and this chiefly because the Christians of these countries belong to the Greek or Eastern Church, though it is difficult to see why this should restrain them from falling in with a change which has many conveniences, and would bring their dates into uniformity with those of the Latin, Teutonic, and Scandinavian nations — an object of increasing importance, as intercommunication is constantly becoming more frequent.

From *Nature* 24 June 1897.

50 YEARS AGO

Apart from its interest to Londoners and its inherent æsthetic merit, the statue of Eros, surmounting the Shaftesbury Memorial Fountain in Piccadilly Circus, London, has a scientific aspect. Originally unveiled in 1893, it is cast in aluminium alloy, the basal structure being of bronze. Its restoration, after a war-time exile, necessitated a thorough cleaning, the repair of a few defects, and the provision of a new bow and arrow. ... The cleaning was carried out with neutral soap and warm water, the surprising discovery being that there is no trace of corrosion arising from the past fifty years in the London atmosphere; with suitable care the original oxide film of the aluminium and the patina of the bronze have both been retained. ... The experience with the restoration of Eros should give an impetus to designers to make more use of these 'modern metals'.

From *Nature* 28 June 1947.

* Workshop on the Latest Developments in Gamma-Ray Bursts, Elba, Italy, 26–27 May 1997.

followed by a fading optical afterglow from a source very near the centre of an apparently distant galaxy⁸, a result so exciting that soon afterwards the workshop was organized by F. Pacini (Arcetri Astrophys. Obs., Florence) and colleagues to review the new multi-wavelength evidence and its implications for GRB models. And, as if to reward the participants, nature provided the most exciting burst so far on 8 May, barely two weeks before the workshop, but time enough to identify an afterglow in X-ray, optical and, for the first time, radio bands.

The most important new evidence presented at this meeting was of redshifted absorption-line features in the optical spectrum³, at wavelengths corresponding to absorption between the burst source and us by iron and magnesium, at a distance of around five billion parsecs. Then this burst came from an explosive emission of about 10^{51} erg of γ -rays (approximately the total radiation from our Sun during the 10^{10} -year age of the Universe) near the edge of the Universe. There is still a small possibility that the fading optical source is not associated with the γ -ray burst, but for most of us it has finally ended the distance-scale debate of the past three decades.

The huge GRB-source luminosities imply a very rapid conversion of energy, equivalent to 0.1 per cent of the total rest-mass energy of a solar-mass star, into γ -rays and afterglow. Observed time structures of γ -ray intensities indicate that, initially, the sources are typically not more than 100 km across (ten times the radius of a neutron star). The merger of two collapsed stars in a binary (for example, two neutron stars or a neutron star and a black hole) is the most popular proposed model for a cosmological burst source.

The enormous explosive local energy release would then initiate a relativistically expanding fireball whose properties and evolution are not sensitive to details of the explosion that caused it. Almost all of the initial fireball energy is eventually converted into a blast whose energy is carried mainly by its relativistically expanding nucleon flow. It is the later interaction of the relativistic blast wave with the ambient medium around the exploded source that is thought to be the origin of the observed X-ray, optical and radio afterglows. Progressively lower-energy spectra are emitted as the blast is slowed by this interaction. There was a consensus among modellers at the workshop that, to account for what is observed in the optical and X-ray afterglows, the Lorentz factor of the blast before such slowdown must be between 100 and 1,000 (that is, the blast's energy was 100 to 1,000 times its nucleon rest-mass energy). To achieve blast-wave Lorentz factors exceeding 100, no more than 10^{-5} of a solar mass of nucleons could have been included in an initial 10^{51} -erg fireball. The observed afterglows would be expected when the blasts expand to a radius exceeding 10^{16} cm, the precise value depend-

ing on the assumed ambient density.

Radio emission (with a peak flux at 8.46 GHz), first detected five days after the 8 May burst⁹ and flaring to twice its initially detected magnitude in one day (S. Kulkarni, Caltech), gives a special clue for just such a blast wave: a strong, incoherent radio flux implies an emitting region at least 10^{16} cm across to avoid the intensity limitation from re-absorption of the radio waves by the electrons that emitted them.

Some were still unpersuaded about the cosmological distances. Particularly controversial¹⁰ was a report that the optical afterglow associated with the 28 February GRB moved by about 1.5 milliseconds of arc per day¹¹ — very hard for cosmological models to explain (T. Bulik, Copernicus Center). Unfortunately, this source is now behind the Sun and we must wait until August to resolve this question. (No proper motion was observed for the very precise radio-location of the 8 May GRB radio afterglow.) And there are claims that the nearby nebula — the supposed galaxy host to the burst — may have changed in either size (D. Lamb, Univ. Chicago) or brightness.

If the relativistic-fireball model is established, GRB afterglows will become a tool to explore properties of the ambient gas in the early Universe. Perhaps, when enough becomes known about optical-afterglow absorption line redshifts and GRB intensities, these events may also yield the necessary relationship between redshift and distance to establish crucial cosmological parameters such as the density of the Universe, or even the cosmological constant (T. Piran, The Hebrew Univ.).

More BeppoSAX designated positions will come in roughly once a month, and we expect to see afterglows from many of them. Afterglows may also be identified in some cases where the GRB is not, perhaps because the γ -rays are mainly beamed away from us. It is hard to point to any other such important astrophysical problem that should be so quickly solved as the origin of γ -ray bursts. But that will probably only be the first extraordinary chapter of a book that promises to become even more exciting. □

Joshua S. Bloom is at the Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK. Malvin Ruderman is in the Department of Physics, Columbia University, New York, New York 10027, USA.

1. Fishman, G. J. & Meegan, C. *Annu. Rev. Astron. Astrophys.* **33**, 415–458 (1995).
2. Djorgovski, S. G. *et al. Nature* **387**, 876–878 (1997).
3. Metzger, M. R. *et al. Nature* **387**, 878–880 (1997).
4. Meegan, C. *et al. Nature* **355**, 143–145 (1992).
5. Paczyński, B. *Publ. Astron. Soc. Pacif.* **107**, 1167–1175 (1995).
6. Lamb, D. Q. *Publ. Astron. Soc. Pacif.* **107**, 1152–1166 (1995).
7. Costa, E. *et al. IAU Circ.* Nos 6572 & 6576 (1997).
8. van Paradijs, J. *et al. Nature* **386**, 686–689 (1997).
9. Frail, D. A. *et al. IAU Circ.* No. 6662 (1997).
10. Sahu, K. C. *et al. Nature* **387**, 476–478 (1997).
11. Caraveo, P. A., Mignani, R., Tavani, M. & Bignami, G. F. *IAU Circ.* No. 6629 (1997).

Daedalus

A level viewing field

A video signal takes up a vast bandwidth. Many coding systems seek to reduce it; some exploit the properties of fractals. Daedalus has an extreme case.

A planar fractal such as the Julia set is generated from a simple mathematical seed, but contains unlimited detail on all scales. A video of such details could be neatly transmitted merely by giving the location of successive frames within the fractal. A suitably complex fractal should contain all possible pictures; indeed, feels Daedalus, there should be an infinite number of such fractals. Just as some fractal forms are particularly good at generating geographical landscapes, others should be good at generating typical TV pictures.

So DREADCO mathematicians are now hunting for the ideal TV-generating fractal. Their goal is a coding process which locates each frame within the fractal, and merely sends its address. The camera and receiver will need massive computing power, but mere telephone bandwidths could link them. And where is massive computing power linked by telephone lines? On the Internet, that's where.

At present, television is a 'top-down' oligarchy. There are few stations, all very rich, and millions of receivers, most quite poor. TV is thus a way of spreading the dreams of the rich. In a thousand subtle ways it invites its audience to yearn for an impossibly rich, chic, exciting lifestyle. Millions of viewers base their human ideals on a few 'celebrities' — a celebrity being someone on whom a TV station is prepared to lavish its vast air-time costs. But when efficient video coding allows the Internet to become the dominant TV provider, the whole notion of a TV station or channel or celebrity will vanish. Just as anybody can set up a Web site, so anyone will be able to launch a TV programme. The result will be a much more democratic TV network, 'convivial' in Illich's sense.

The consequences could be profound. The rigidly stylized TV dream-dictatorship, shaped by big-company advertising, will be replaced by a chaotic uncensorable mass of video images, shaped by individual greed, egotism and missionary zeal. It will be wildly interactive. Any viewer could reply to any programme, or send in real-time suggestions or additions, or launch a rebuttal or diversion. What video culture will ultimately crystallize from the chaos, Daedalus cannot guess, any more than he can guess where the Internet itself is going. But the result must be healthier than what we have now.

David Jones

A peptide antibiotic from human skin

To avoid opportunistic infections, plants and animals have developed antimicrobial peptides in their epithelia that can form pores in the cytoplasmic membrane of microorganisms¹. After contact with microorganisms, vertebrate skin², trachea and tongue epithelia³ are rich sources of peptide antibiotics¹, which may explain the unexpected resistance of these tissues to infection. Here we report that human skin is protected in a similar way by an inducible, transcriptionally regulated, antibiotic peptide, which resembles those in other mammals.

Patients with psoriasis have fewer skin infections than expected⁴, leading us to speculate that lesional psoriatic skin might produce antimicrobial peptides. Cationic antimicrobial peptides bind strongly to their target bacteria, so we isolated and purified peptide antibiotics from psoriatic scale extracts using a whole *Escherichia coli* affinity column. We subsequently purified bound peptides, which demonstrated antimicrobial activity in a plate assay, to homogeneity using high-performance liquid chromatography. We recovered 200–400 µg of a pure antimicrobial peptide (relative molecular mass 4,000) from 50-g samples of psoriatic scales. Amino-acid sequence analysis of this peptide (Fig. 1) revealed the consensus sequence of a β-defensin with homology to bovine tracheal and lingual antimicrobial β-defensins³, as well as human β-defensin-1 (hBD-1)⁵.

Using degenerate primers based on amino-acid sequence data, we generated the complete complementary DNA sequence from RNA obtained from human foreskin-derived primary keratinocytes. The deduced amino-acid sequence of the precursor peptide (Fig. 1) consisted of 41 residues present in the mature peptide as well as a leader sequence, which may indicate that this is a secreted peptide. The highest homology is to bovine tongue- and trachea-derived antimicrobial peptides³, produced in cow epithelia in response to microorganisms or inflammatory cytokines. We conclude that the peptide, named hBD-2, is the second human β-defensin to be found. The first human β-defensin (hBD-1) is mainly produced by epithelia from the urogenital tract and, to a lesser degree, in trachea and lung⁶, and was originally discovered in human blood filtrates⁵.

We found that hBD-2 was highly effective in killing Gram-negative bacteria (*E. coli*, *Pseudomonas aeruginosa*) with LD₉₀ (the dose that achieves 90% reduction of colony-forming units) near 10 µg ml⁻¹. The yeast *Candida albicans* was also effec-

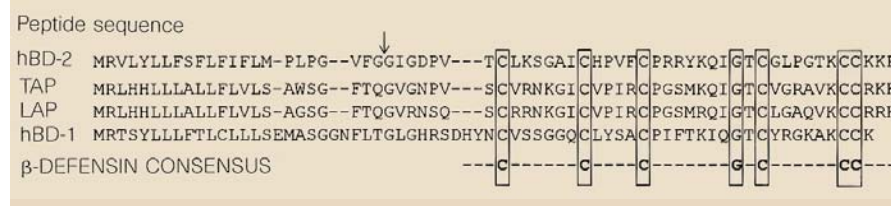


Figure 1 The deduced amino-acid sequence (single-letter code) of the hBD-2 precursor based on the complementary DNA sequence obtained from human keratinocytes, with the sequences of bovine tracheal antimicrobial peptide (TAP), bovine lingual antimicrobial peptide (LAP) and human β-defensin-1 (hBD-1), and the β-defensin consensus sequence. Arrow indicates the amino terminus of hBD-2 obtained from psoriatic scales. Isolation of the peptide was as described in ref. 9. The full-length cDNA structure was derived using an inverse PCR method¹⁰ using specific primers according to the sequence found with degenerate primers. The cDNA sequence is available from the EMBL/Genbank database, accession number Z71389.

tively killed (LD₉₀, 25 µg ml⁻¹), but hBD-2 achieved a bacteriostatic effect on the Gram-positive *Staphylococcus aureus* only at concentrations as high as 100 µg ml⁻¹.

Using reverse transcription and semi-quantitative polymerase chain reaction (RT-PCR) we saw low hBD-2 messenger RNA expression in foreskin-derived keratinocytes. Expression was greatly upregulated with tumour-necrosis factor-α within 1 h of stimulation and persisted for more than 48 h. Gram-negative and Gram-positive bacteria, as well as *C. albicans*, strongly induce hBD-2 (Fig. 2), in contrast

to hBD-1, which is not upregulated by inflammatory stimuli⁶. Thus hBD-2 is the first human β-defensin found to be regulated at a transcriptional level in response to contact with microorganisms.

We detected constitutive hBD-2 expression in freshly isolated foreskin, lung and trachea (Fig. 2). In contrast, considerably less hBD-2 is expressed in kidney, uterus and salivary gland tissue. Both small intestine and liver mRNA preparations failed to demonstrate hBD-2 expression. Therefore, hBD-2 may represent the human counterpart of the bovine lingual and tracheal peptides.

Our observations show that human skin contains a chemical shield of endogenous peptide antibiotics, which are induced by contact with microorganisms and are operative in other human epithelia. Disruption of this shield, as in cystic fibrosis^{7,8}, might be a reason for recurrent infections of skin and other epithelia. It is intriguing to speculate that human peptide antibiotics, like hBD-2, might be ideal therapeutic agents, avoiding the problems of acquired resistance.

J. Harder, J. Bartels

E. Christophers

J.-M. Schröder

Department of Dermatology,

University of Kiel,

24105 Kiel, Germany

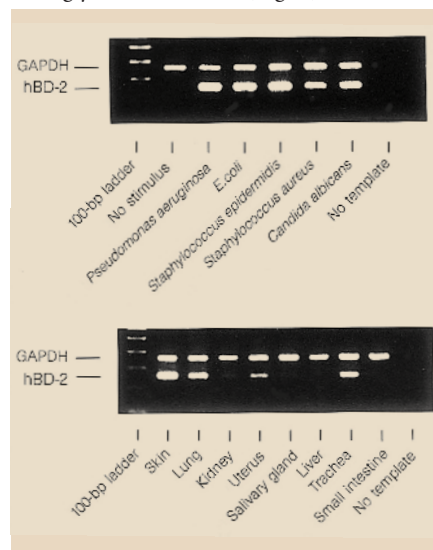


Figure 2 hBD-2 mRNA is upregulated by contact with different microorganisms (upper panel) and constitutively expressed in organs other than skin (lower panel). Foreskin-derived, third-passage keratinocytes were stimulated with different heat-killed microorganisms (10⁷ per ml) for 16 h and analysed for mRNA expression by semi-quantitative RT-PCR using hBD-2 primers (5'-CCAGCCATCAGCCA-TGAGGGT-3'; 5'-GGAGCCCTTCTCGAATCCGCA-3') and primers for glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) as described¹⁰. Controls: 'no stimulus', buffer without bacteria; 'no template', primer only, no cDNA.

1. Boman, H. G. *Annu. Rev. Immunol.* **13**, 61–92 (1995).
2. Zasloff, M. *Proc. Natl Acad. Sci.* **84**, 5449–5453 (1987).
3. Schonwetter, B. S., Stolzenberg, E. D. & Zasloff, M. A. *Science* **267**, 1645–1648 (1995).
4. Henseler, T. & Christophers, E. *J. Am. Acad. Dermatol.* **32**, 982–986 (1995).
5. Bensch, K. W., Raida, M., Mägert, H.-J., Schulz-Knappe, P. & Forstmann, W.-G. *FEBS Lett.* **368**, 331–335 (1995).
6. Zhao, C., Wang, I. & Lehrer, R. I. *FEBS Lett.* **396**, 319–322 (1996).
7. Smith, J. J., Travis, S. M., Greenberg, E. P. & Welsh, M. J. *Cell* **85**, 229–236 (1996).
8. Goldman, M. J. *et al. Cell* **88**, 553–560 (1997).
9. Schröder, J.-M., Gregory, H., Young, J. & Christophers, E. J. *Invest. Dermatol.* **98**, 241–247 (1992).
10. Bartels, J. *et al. Biochem. Biophys. Res. Comm.* **225**, 1045–1051 (1996).

The zebrafish organizer requires *chordin*

The dorsoventral pattern of vertebrate embryos is established and regulated by opposing gradients of ventralizing bone morphogenetic proteins (BMPs) and BMP antagonists¹⁻³ (such as Chordin^{2,4}, Noggin^{5,6} and Follistatin^{7,8} in the frog embryo) secreted by a dorsal organizer. In the zebrafish, *Danio rerio*, mutations in a number of genes that affect dorsoventral patterning of the early embryo have been identified^{9,10}, but only two, *dino* and *mercedes*, are involved in dorsal specification¹¹, with *dino* mutants displaying the stronger, ventralized phenotype. Here we show that

the *dino* phenotype is caused by a mutation in the zebrafish *chordin* gene, revealing that Chordin is an essential component of the dorsal organizer. Furthermore, we provide evidence that Chordin and BMPs are antagonistic, not only at the protein level but also at the transcriptional level.

Zebrafish Chordin (S. S.-M. & M. H., manuscript in preparation) shows the same protein organization as its *Xenopus* and *Drosophila* homologues, Xchd² and Sog¹²⁻¹⁴, respectively (Fig. 1a). A hydrophobic amino-terminal sequence is followed by the first of four cysteine-rich repeats (CR1); the remaining three (CR2-4) being towards the carboxy terminus. The *chordin* gene is first expressed on the dorsal side of the embryo, starting before gastrulation (Fig. 2a), when the Spemann organizer is thought to generate dorsalizing signals.

In a restriction-fragment length polymorphism (RFLP) linkage analysis of more than 400 mutant chromosomes, we saw no recombination between *dino* and *chordin* (Fig. 1b), indicating that *dino* and *chordin* are within 0.25 centimorgans (cM). The *chordin* complementary DNA from *din*^{tt250} mutant embryos has a deletion of 104 base pairs, causing a frame-shift predicted to produce a severely truncated protein (Fig. 1c). This truncated form has only the first 42 amino acids of the wild-type Chordin protein and completely lacks the cysteine-rich repeats.

Injection of *Xenopus chordin* messenger RNA into *dino* mutants rescues the mutant phenotype, producing viable and fertile homozygous *dino* mutant adult fish (Fig. 1d,e,f). This indicates that *chordin* is not required after the early stages of development, as the injected RNA is largely degraded by the end of the gastrulation stage (data not shown). These data indicate that the defects of *dino* mutant embryos are potentially caused by a null mutation in the *chordin* gene. In view of this genetic relationship, we will refer to both the mutant and the gene as *chordin*.

In mutant embryos, *chordin* expression is strongly reduced at midgastrula stages (Fig. 2a,b), indicating that *chordin* expression requires functional Chordin protein. Indeed, the wild-type *chordin* expression pattern can be rescued in *chordin* mutants by the injection of both *Xenopus chordin* mRNA (Fig. 2c) and mRNA encoding a dominant negative BMP receptor (Fig. 2d), whereas over-expression of *bmp-4* in wild-type embryos leads to a reduction in *chordin* expression levels similar to those in *chordin* mutants (data not shown). This suggests that such autoregulation of *chordin* is achieved by an inhibition of antago-

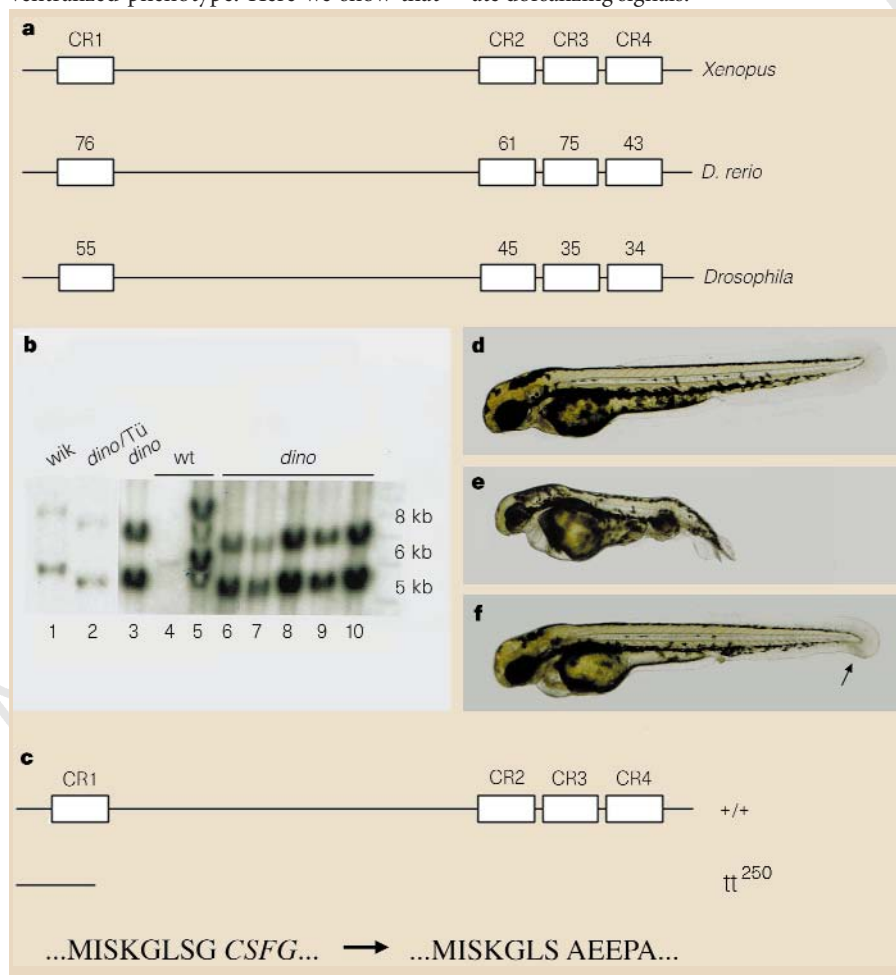


Figure 1 **a**, Schematic representation of *Xenopus* Chordin, zebrafish Chordin, and *Drosophila* Sog. Numbers indicate the percentage of amino-acid identity between the respective cysteine-rich domains. **b**, Linkage analysis, carried out as in ref. 15. Genomic DNA was digested with EcoRI. Lanes 1 and 2: parental DNA from the WVK strain and a heterozygous *dino* individual in the Tübingen background, respectively. Lanes 3, 6-10: DNA from pools of *dino* mutant F₂ embryos (a total of 136 genome equivalents; 272 mutant chromosomes). Lanes 4 and 5: DNA from pools of sibling F₂ embryos (+/+ and *dino*/+; 0.5 and 24 genome equivalents, respectively). **c**, Predicted structure of *din*^{tt250}, containing the first 42 wild-type amino acids and 54 additional unrelated residues. Part of the amino-acid sequence of the wild type (left) and mutant protein (right) is shown, with the first four amino acids of CR1 displayed in italics. **d**, Wild-type embryo at 3 days of development. **e**, Sibling *dino* mutant embryo. Note the small eyes and split ventral tail fin. **f**, Mutant embryo, rescued by injection of *Xenopus chordin* mRNA. The ventral tail fin is partially duplicated (arrow), but otherwise the embryo appears normal.

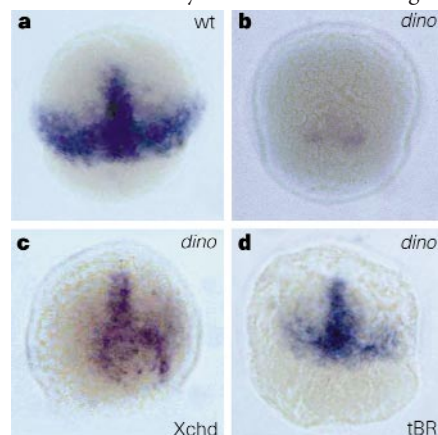


Figure 2 **a**, Wild-type and **b**, Mutant embryo. *chordin* mRNA can be rescued to almost wild-type levels by the injection of mRNA encoding **c**, *Xenopus chordin* (Xchd) or **d**, A dominant-negative BMP receptor (tBr). All embryos are shown in a dorsal view after *in situ* hybridizations with a *chordin* anti-sense probe. **c**, **d**, 70% epiboly. **a**, **b**, 80% epiboly. RNA injections were done as in ref. 3.

nizing BMP signals which normally act as negative regulators of *chordino* expression.

In addition, *chordino* mutants display a dorsally expanded *bmp-4* expression domain³, indicating that during normal development *chordino* has a negative influence on *bmp-4* transcription in dorsolateral regions. Again, this negative regulation might be mediated by BMPs themselves, which normally act as positive regulators of their own expression³.

It has been shown that *Xenopus* Chordin inhibits the antagonizing Bmp-4 activity by direct binding to the Bmp-4 protein, thereby preventing receptor activation⁴. Our data demonstrate that this inhibition at the protein level is potentiated by two synergistic feedback loops which lead to a derepression of *chordino* transcription and an inactivation of *bmp-4* expression³.

Our findings demonstrate that *chordino* is required during early dorsoventral patterning of the zebrafish embryo, where it functions as an antagonist of ventralizing BMP signals. This supports the notion that mechanisms of early dorsoventral patterning are conserved between vertebrates and invertebrates¹³.

Stefan Schulte-Merker

MPI für Entwicklungsbiologie, Spemannstrasse 35,
72076 Tübingen, Germany

Kevin J. Lee

Howard Hughes Medical Institute,
Department of Biochemistry and Molecular
Biophysics,
Center for Neurobiology and Behavior,
Columbia University, New York,
New York 10032, USA

Andrew P. McMahon

Department of Molecular and Cellular Biology,
Harvard University, 16 Divinity Avenue,
Cambridge, Massachusetts 02138, USA

Matthias Hammerschmidt*

MPI für Immunbiologie, Stübeweg 51,
79108 Freiburg, Germany
e-mail: hammerschmidt@immunbio.mpg.de

* To whom correspondence should be addressed.

Have quantum scars been observed?

Wilkinson *et al.*¹ have reported observations of oscillations in the current of a resonant tunnelling diode in a magnetic field. They proposed that this provides the first direct experimental evidence for quantum 'scarring' in a real chaotic system. But we believe that Wilkinson *et al.* have not observed quantum scars in the usual sense. They considered a far-from semiclassical regime where current theories of scarring are not valid. Demonstrating these effects in a solid-state device is a tremendous achievement. However, we believe that even in the semiclassical limit this type of experiment does not go beyond other experiments on atoms in magnetic fields as a probe of scarring.

A quantum particle travelling from A to B experiences all possible paths in between. The wave-like nature of a quantum particle means that each path has a phase and so neighbouring paths interfere with each other. The combined contributions of these paths build up the wavefunction, which provides a complete description of the quantum dynamics of the particle. At high energies, in the semiclassical limit ($\hbar \rightarrow 0$), phases associated with different paths are large. Hence, neighbouring paths can rapidly fall out of phase and their contributions cancel by destructive interference. When the typical motion is a chaotic trajectory, this should produce wavefunctions with no structure other than a random grainy pattern.

But Heller² found that for $\hbar \rightarrow 0$, quantum states of a chaotic system show 'scars', concentrations near the paths of unstable, isolated periodic orbits (classical orbits which retrace themselves and are isolated if there is no other periodic orbit of similar phase nearby). The issue of how and when quantum states are scarred in the semi-

classical limit has sparked a lively debate. The mathematical theories all assume that the periodic orbits are isolated, though more than one isolated orbit can scar a single state.

Wilkinson *et al.* have probed the 25 or so lowest quantum levels above the ground state. By lowering the energy, the phases of the classical orbits are made small, so that they remain in-phase. Hence, the neighbourhoods of topologically distinct periodic orbits become so large that they overlap and so cease to be isolated.

In the experiment, the electronic motion is confined by a magnetic field to a cylindrical energy surface whose ends are the two barriers of the well. Wilkinson *et al.* found that the tunnelling is dominated by four single eigenstates, roughly 'linear' in shape with $N \approx 10$ oscillations along the magnetic field. However, the energy surface supported only $n \approx 2-3$ oscillations perpendicular to the field, so each 'linear' eigenstate is supported by a large fraction of the Poincaré surface on the emitter wall, including a cluster of short periodic orbits of similar action that traverse the length of the cylinder.

This similarity of action is analogous to back and forth motion close to the axis of a long, thin cylinder. Most of the action comes from the longitudinal motion, with smaller differences due to the transverse components. The periodic orbits become isolated only when we can resolve the differences in action (typically for $n > 10$). For small n (non-isolated case), sequences of individual 'linear' states, roughly aligned with the magnetic field, are ubiquitous^{1,3}. These states cannot distinguish between the different periodic orbits. The stability parameters of the classical dynamics are essentially irrelevant, in contrast to an isolated scar. They are essentially a quantum phenomenon, the so-called adiabatic separability⁴, more relevant in the deep quantum regime than an interpretation in terms of scars. A linear state ($n \approx 1$) is shown in Fig. 1a, and Fig. 1b shows a scar in the semiclassical regime ($n \approx 30$).

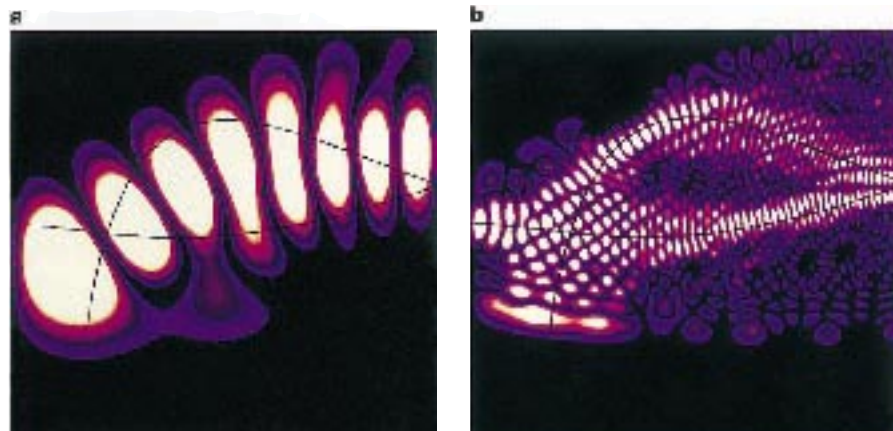


Figure 1 Quantum eigenstates which dominate tunnelling for: **a**, $n \approx 1$ far from the semiclassical limit showing 'linear' quantization; and **b**, $n \approx 30$ showing scarring by the S_1 unstable periodic orbit observed in wide-well experiments. We ensured that both states corresponded to exactly the same set of classical periodic orbits, differing only in effective size of $\hbar$ by using a scaling property of the dynamics.

- Graff, J. M., Thies, R. S., Song, J. J., Celeste, A. J. & Melton, D. A. *Cell* **79**, 169–179 (1994).
- Sasai, Y., Lu, B., Steinbeisser, H. & De Robertis, E. M. *Nature* **376**, 333–336 (1995).
- Hammerschmidt, M., Serbedzija, G. N. & McMahon, A. P. *Genes Dev.* **10**, 2452–2461 (1996).
- Piccolo, S. Y., Sasai, Y., Lu, B. & De Robertis, E. M. *Cell* **85**, 589–598 (1996).
- Re'em-Kalma, Y., Lamb, T. & Frank, D. *Proc. Natl Acad. Sci. USA* **92**, 12141–12145 (1995).
- Zimmerman, L. B. & Harland, R. M. *Cell* **85**, 599–606 (1996).
- Hemmati-Brivanlou, A., Kelly, O. G. & Melton, D. A. *Cell* **77**, 283–295 (1994).
- Fainsod, A. *et al. Mech. Dev.* (in the press).
- Haffter, P. *et al. Development* **123**, 1–36 (1996).
- Driever, W. *et al. Development* **123**, 37–46 (1996).
- Hammerschmidt, M. *et al. Development* **123**, 95–102 (1996).
- Francois, V., Solloway, M., O'Neill, J. W., Emery, J. & Bier, E. *Genes Dev.* **8**, 2602–2616 (1994).
- Holley, S. A. *et al. Nature* **376**, 249–253 (1995).
- Biehls, B., Francois, V. & Bier, E. *Genes Dev.* **10**, 2922–2934 (1996).
- Schulte-Merker, S., Eeden, F. V., Halpern, M. E., Kimmel, C. B. & Nüsslein-Volhard, C. *Development* **120**, 1009–1015 (1994).

Periodic-orbit effects have been extensively investigated using highly excited atoms in magnetic fields. The classic 1969 experiment⁵ revealed a set of evenly spaced modulations in atomic photoabsorption spectra near the ionization threshold, associated with quantum states scarred by an unstable periodic orbit.

T. S. Monteiro

Department of Physics and Astronomy,
University College London,

Gower Street, London WC1E 6BT, UK

D. Delande

Laboratoire Kastler Brossel, 4 Place Jussieu, Tour 12,
Premier Etage, F-75252 Paris Cedex 05, France

J.-P. Connerade

Blackett Laboratory, Imperial College,
Prince Consort Road, London SW7 2BZ, UK

Fromhold et al. reply — Unstable but periodic classical orbits are of fundamental importance in quantum chaos^{6,7}. They produce regular clustering of the energy levels⁶ and certain states exhibit regions of enhanced probability density or ‘scars’ in the neighbourhood of the classical periodic trajectories². The scarring effect seems to have been unexpected in view of the instability of the periodic orbits⁸. The typical trajectory is chaotic and irregular, which would be expected to correspond to wavefunctions exhibiting irregular and diffuse patterns of probability density. We recently showed¹ that scarred and unscarred states can make very different contributions to a physically measurable quantity, in our case the tunnel current through a semiconductor device.

There is much published work on the photoabsorption spectrum of highly excited hydrogenic atoms in a strong magnetic field, which show modulations due to the effects

of unstable periodic orbits in the classically chaotic regime. Owing to the complexity of the spectra at high excitation energies in the semiclassical limit, this has provided only indirect evidence for the scars confined to individual eigenstates^{9–11}. The 1969 experiment⁵ referred to by Monteiro *et al.* gave a beautiful demonstration of the role of unstable periodic orbits in quantum chaos, but had insufficient resolution to distinguish between the effects of wavefunction scarring and energy-level clustering on the photoabsorption spectra. However, Müller and Wintgen¹² noted that the scarring phenomenon survives the semiclassical limit and found that, at low excitation energies, wavefunctions were scarred by short-period orbits. It is this regime that our resonant tunnelling experiments were designed to investigate¹.

Our theoretical studies of the electronic levels in a 120-nm-wide quantum well of a resonant tunnelling diode in the chaotic regime revealed subsets of states, strongly scarred by particular periodic orbits¹³, and also indicated that the current through the device was dominated by tunnelling via the scarred states in the well. But our experiments on the 120-nm-wide quantum well¹⁴ could not resolve discrete levels because of broadening by scattering processes.

Our work on a 22-nm-wide quantum well¹ with clearly resolved energy levels showed that peaks in the current–voltage characteristic of the device corresponded to electrons tunnelling into the scarred states of the quantum well. This was related directly to the concentration of the probability density of these states close to a cluster of three very similar unstable scarring orbits with almost identical periods. By contrast, the unscarred wavefunctions of adjacent levels have irregu-

lar antinode patterns, which inhibit the tunnel process (Fig. 2). Further, the energies of the scarred states and their relation to orbital period were given accurately by the semiclassical quantization rule^{4,15}. The results are from exact numerical calculations, valid for all energies, and are independent of approximate general theories of wavefunction scarring, which break down in the quantum limit. Note also that the quantum well exhibits particularly strong scarring effects because of its special dynamical properties, such as the absence of spatially closed but aperiodic orbits, which are not shared by atoms.

Monteiro *et al.* argue that our results are not attributable to “scars in the usual sense”, as they do not distinguish between the different short-period orbits of the cluster. But this view is not supported by the literature or generally accepted. Semiclassical scarring of subsets of states by clusters of similar orbits has been shown⁴. For non-isolated orbits, Heller² refers to super-scars resulting from the overlap of many scars and comments that “increasing the density of similar periodic orbits can only enhance the scarring effect”.

The scarring effect has also been studied for the classical electromagnetic eigenmodes of a microwave cavity¹⁶. Here the scarring orbits are the periodic trajectories of rays reflecting around inside the cavity. In this case it was observed that the association of wavefunctions with periodic ray orbits emerges even at low frequencies when the electromagnetic wavelength is only roughly one-quarter of the cavity dimension. Analogous conditions hold in our resonant tunnelling diode where the antinode quantum numbers, (ν) for the observed scarred states ($\nu \approx 6$ –9) are relatively small, as the de Broglie wavelength of the electrons is roughly one-quarter of the width of the quantum well.

But this in no way invalidates the physical concept of scarring as an imprint of an unstable periodic orbit on the quantum state of a classically chaotic system. The doubts raised by Monteiro *et al.* originate from their rather restricted view of scarring and are not substantiated in the literature.

T. M. Fromhold, F. W. Sheard

L. Eaves, P. B. Wilkinson

Department of Physics, University of Nottingham,
Nottingham NG7 2RD, UK

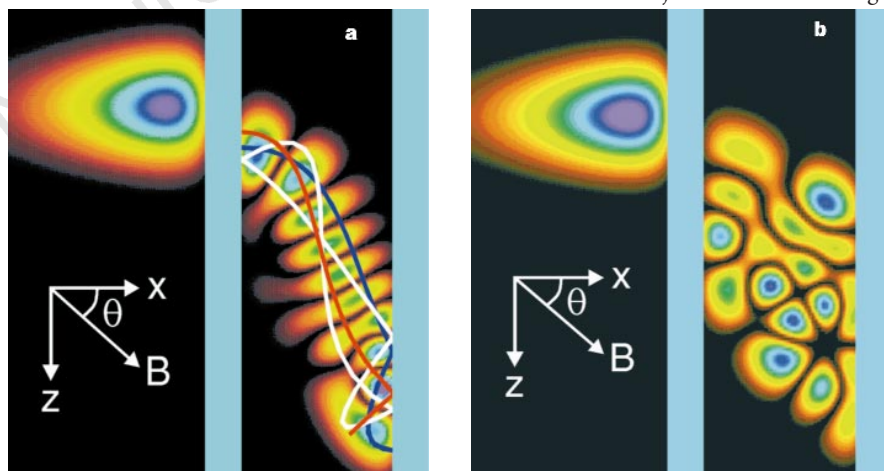


Figure 2 Probability density plots for two neighbouring eigenfunctions of the 22-nm-wide quantum well described in ref. 1. Probability is represented by a linear colour scale with $P=0$ in black, increasing linearly from ‘warm’ to ‘cold’ colours. The probability density of a localized state in the emitter accumulation layer is also shown. Blue rectangles indicate tunnel barriers. The plots are drawn in the x - z plane, where the x -axis is perpendicular to the confining barriers. A magnetic field of 37 T is applied in the x - z plane at an angle $\theta=40^\circ$ to the x -axis. **a**, One of the scarred states identified in ref. 1 with energy $E_n=309.25$ meV. The three scarring orbits are shown overlaid. **b**, The probability distribution for the next highest energy level with $E_{n+1}=315.5$ meV reveals no trace of scarring. A bias voltage of 668 mV was applied across the resonant tunnelling diode containing the well and the effects of conduction-band non-parabolicity were included in the calculation¹.

1. Wilkinson, P. B. *et al.* *Nature* **380**, 608–610 (1996).
2. Heller, E. J. *Phys. Rev. Lett.* **53**, 1515–1518 (1984).
3. Fromhold, T. M. *et al.* *Solid State Electron.* **40**, 7–14 (1996).
4. Eckhardt, B., Høse, G. & Pollak, E. *Phys. Rev. A* **39**, 3776–3793 (1989).
5. Garton, W. R. S. & Tomkins, F. S. *Astrophys. J.* **158**, 839–845 (1969).
6. Berry, M. V. *Proc. R. Soc. Lond. A* **413**, 183–198 (1987).
7. Gutzwiller, M. C. *Chaos in Classical and Quantum Mechanics* (Springer, New York, 1990).
8. Bogomolny, E. B. *Physica D* **31**, 169–189 (1988).
9. Holle, A. *et al.* *Phys. Rev. Lett.* **61**, 161–164 (1988).
10. Wintgen, D. & Hönig, A. *Phys. Rev. Lett.* **63**, 1467–1470 (1989).
11. Heller, E. J. *Nature* **380**, 583–584 (1996).
12. Müller, K. & Wintgen, D. *J. Phys. B* **27**, 2693–2718 (1994).
13. Fromhold, T. M. *et al.* *Phys. Rev. Lett.* **75**, 1142–1144 (1995).
14. Fromhold, T. M. *et al.* *Phys. Rev. Lett.* **72**, 2608–2611 (1994).
15. de Polavieja, G. G. *et al.* *Phys. Rev. Lett.* **73**, 1613–1616 (1994).
16. Sridhar, S. & Heller, E. J. *Phys. Rev. A* **46**, R1728–R1731 (1992).

A richly varied life in physics

A Tale of Two Continents: A Physicist's Life in a Turbulent World

by Abraham Pais

Princeton University Press: 1997. Pp. 511.

\$35, £25

H. B. G. Casimir

Abraham Pais is a well-known figure among today's senior physicists. In the rather esoteric group of theorists dealing with particles and fields, he ranks as one of the foremost experts, especially on so-called weak interactions. But he is also an expert in a broader sense, as is borne out by his book *Inward Bound: Of Matter and Forces in the Physical World* (Oxford University Press, 1986). This was preceded by "*Subtle is the Lord*" (OUP, 1982), a widely acclaimed biography of Albert Einstein, and was followed by the biography *Niels Bohr's Times: In Physics, Philosophy and Polity* (OUP, 1991).

No wonder Pais has been much in demand as a speaker and often called on to help organize scientific meetings. Prospective readers of his autobiography will therefore expect a richly varied story. They will not be disappointed.

In trying to give an idea of the rich contents, I shall not follow strict chronological order (neither does Pais). But five periods of his life can be distinguished: his youth until the end of the Second World War and of the occupation of the Netherlands (1918–45); a brief stay in Copenhagen (1946); a career at the Institute for Advanced Study in Princeton (1946–63); a professorship at Rockefeller University in New York (1963–88); and his emeritus professorship, with a desk in New York and another in Copenhagen (from 1988 onwards).

The beginnings of the first period were uneventful. Abraham — Bram to his friends — is the son of a Jewish schoolteacher father, belonging to the Portuguese-Jewish (or Sephardic) community, whose ancestors settled in Amsterdam centuries ago. His mother was a teacher as well and, although the family was not wealthy, it managed to get by. Bram was a bright pupil, achieving high grades with little effort. While at home he adhered to the Jewish rites. He soon lost his faith, but became interested in Zionism. Then came the war, Germany's occupation of the Netherlands and the regime's stringent anti-Jewish measures.

Pais just managed to obtain his doctorate before going into hiding. He was found but escaped by the skin of his teeth, thanks to the help of a plucky young woman who, armed with a letter from the German Nobel laureate Werner Heisenberg, went to plead with a high-ranking Nazi official. All this took place against the background of an ever

worsening famine. Bram's parents escaped too. His only sister and her husband refused to look for a hiding place, were deported and did not return. The same fate befell George Uhlenbeck's former assistant, Boris Kahn.

Throughout his period in hiding, Pais continued to do theoretical work, some of which was later published. As far as I know, it had no lasting influence, but it must have been a great help in keeping up his morale. One cannot but admire the way in which the young Pais kept up his spirits and how he now writes about his experience.

After the war, Pais's stay in Copenhagen lasted less than a year, during which he acted as a kind of private secretary to Niels Bohr, helping to prepare lectures. The friendships that Pais established in Denmark were confirmed and extended on many later visits.

Pais writes of Bohr: "No other physicist, for that matter no one else, ever has stirred in me feelings of such high respect and deep affection. He is my father figure. Only later did I come to understand how fortunate I have been to spend a period with Bohr that was neither too short nor too long. Why not too long? Because it can be dangerous for a young man in his formative years to spend too much time in the presence of a personality as powerful as Bohr's." However that may be, if Pais, following Einstein, ever had any misgivings about the interpretation of quantum mechanics, they disappeared. He became a staunch supporter of Bohr's ideas on phenomena and complementarity.

In the autumn of 1946, Pais travelled to the United States and settled in Princeton. Not everyone will agree with his views on the American way of life, but his stories are amusing.

At the Institute for Advanced Study he rose through the ranks, becoming finally a full professor with tenure. Why then did he choose in 1963 to move to New York? He states as his reason that he was afraid of becoming more and more of a manager rather than a scientist. At that time, however, his best work in theoretical physics had been done. At Rockefeller University we see him change gradually into a historian of science.

As to his personal life, Pais has married three times. His first marriage lasted from 1956 to 1962. He is reticent but unembittered about the separation and writes with true affection about his son, born in 1958, and about his daughter-in-law and grandson. He is rather sarcastic about his second marriage which lasted from 1976 to 1985: "her running off was the best thing she had ever done for me". Finally, in 1990, he finds the ideal partner, about whom he writes with loving praise verging on adulation.



Pais: theoretical physicist turned historian.

Pais has done an enormous amount of travelling, as a tourist, as a visiting scientist and, more often, as a combination of the two. His trips have often involved mountaineering: in 1960 he climbed Mont Blanc, and in 1962 the Matterhorn. With modern equipment and fixed ropes, these climbs are less difficult and dangerous than they were a century ago in the days of the pioneering mountaineer Edward Whymper. But both are strenuous and a remarkable feat for any amateur over 40.

Scattered through the text are many anecdotes and incisive remarks about famous scientists. There are also sections on world affairs — where, much to my surprise, Pais mentions the World Series baseball championship — and on the art of writing a book.

But it is obvious that Pais, for all his multifarious interests and abilities, is, or was, primarily a hardworking and enthusiastic theoretical physicist. Time and time again he mentions the problems he has worked on and recalls his students and colleagues. He started out with problems in quantum electrodynamics, which was then reaching the stage at which experimental results could be predicted with surprising accuracy, although Pais never found the theory entirely satisfactory.

The bulk of his work deals with particle physics — an entirely new branch of physics with little contact with or influence on the physics of the first half of our century. Pais

endeavours to explain at least the gist of its development, a development to which he himself made important contributions. Has he succeeded? I am not sure. That the treatment of the mathematics is rather sketchy does not matter for the professional physicist: one can always turn to *Inward Bound* and, if one wants to become a real expert, from there to one of several good textbooks. But has Pais given nonphysicists an inkling of the reaction of physicists to the sensational news of charge parity violation? Can he convince authorities that it is worth spending thousands of millions to create a few extremely short-lived particles? Perhaps that is asking too much.

No book is perfect. Although Pais is on the whole very readable and both instructive and entertaining, he is sometimes a bit long-winded. Several sections, in particular those dealing with travel, could do with pruning, and some duplication might have been avoided. Pais himself recommends an index of subjects, but does not follow his own advice. But these minor shortcomings in no way detract from the value of this impressive book. □

H. B. G. Casimir is at De Zegge 7, 5591 TT Heeze, The Netherlands.

Our flexible friends

Cells, Embryos and Evolution: Towards a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability

by John Gerhart and Marc Kirschner
Blackwell Science: 1997. Pp. 642.
£29.50, \$64.95

J. M. W. Slack

Every biologist has heard of Darwin's finches and knows that the shape of the beak may evolve differently depending on the food source. But how many ways are there to change a beak? Is the range of possible beaks limited by developmental constraints? Are modern birds better at changing their beaks than their remote forebears? These questions are posed in the preface to *Cells, Embryos and Evolution* to emphasize that now is the time to apply our newly found molecular understanding of development to find out what really happened in evolution.

John Gerhart and Marc Kirschner are distinguished experts in cell and developmental biology. As might be expected, they paint a mighty canvas and fill it with good things. Much of their text is devoted to a detailed explanation of the main developmental processes in vertebrates and insects, rather less to some important themes in cell biology, and several chapters to an investigation of how the organization of developmental systems has facilitated



Bear essentials

Conventional wisdom says you should never come between a mother bear and her cubs. Inadvertently, Erwin A. Bauer did just that when he fell asleep on a hiking trail in Grand Teton National Park in Wyoming. But the bears ignored him and he survived to recount the tale in the introduction to his book, *Bears: Behavior, Ecology, Conservation* (Voyageur Press, \$35).

This is the latest in a series of popular books from Voyageur exploring the world of North American bears. In the above picture, one of many fine photos in the book by Erwin and Peggy Bauer, a mother grizzly disciplines a surprised cub that has been wandering too far away. The author gives tips for capturing these elusive creatures on film.

the evolution of morphological novelty.

Their grand theme is that the developmental systems of metazoa show "flexibility and robustness", characteristics that may enable mutations to produce significant changes in morphology without always being disadvantageous. The elements of such a system are things such as parallel pathways, weak rather than strong regulatory connections, or competitive and exploratory mechanisms rather than fully determined ones. There may be a progressive increase in 'evolvability' of this kind over evolutionary time. There is no doubt that the developmental mechanisms of vertebrates and insects, which are the two best understood animal groups, show these features, albeit in rather different ways.

But, if we are to ascribe their success in producing many varied morphologies to such characteristics, then we must ask whether other groups that have failed to produce such variety have a developmental system that is demonstrably more rigid. This could possibly be argued of the nematodes, whose morphological diversity as a phylum does not seem to match that of the arthropods or chordates. Early studies of *Caenorhabditis elegans* did emphasize its rigid cell lineage, but the more the nematode has been studied the less 'hard wiring' of events has been evident and the more signalling processes have been discovered, which would count as "flexible" to Gerhart and Kirschner.

Things are even less clear when we encounter the 'priapulid problem'. The priapulids, named after the eponymous Priapus with his large male generative organ, are an entire phylum containing only 17 species that seem to have changed hardly at all since their appearance in the Middle Cambrian. We have no idea whether they have stayed the same because their developmental systems are inherently less evolvable than more glamorous phyla, or whether there has simply been little selective pressure for diversification because conditions down in the marine mud have not changed much in the past 530 million years.

The book tends to play down the importance of 'constraints', which are the usual concern of developmental biologists. It was at one time supposed that changing almost anything in early development would be impossible because of the knock-on effects on later stages. But this view has tended to moderate as people have become more familiar with the wide range of early adaptations to different reproductive lifestyles. It is now recognized that the stage of maximum similarity between members of a taxon is not the earliest embryonic stage, but the middle stage. This is often called the 'phylotypic' stage, and receives many mentions in the book. But the authors are a little coy about committing themselves to the existence of any phylotypic stages other than the familiar pharyngula of the vertebrates and extended germ band of the insects.

They tentatively suggest that a "comaform" nematode might be phylotypic, but neglect other possible candidates such as the veliger larva of gastropods, or an early segmenting stage of an annelid worm. In accordance with their general case they argue that the conservation of the phylotypic stage is due to its flexibility and robustness, and do not seriously consider the alternative view that it has stayed the same because selective pressures are at a minimum in the middle stages of development.

The most spectacular achievement of molecular developmental biology has been the discovery of long-range homologies between animal groups in terms of things such as the molecular basis of anteroposterior and dorsoventral patterning. These discoveries enable us to reconstruct long dead ancestors in a way that would have been unthinkable a few years ago.

Gerhart and Kirschner use these data to argue that the stem metazoan was a "roundish flatworm" that flourished below the surface of the Vendian biot. This organism had Hox genes specifying the anteroposterior pattern; *emx/otx* genes to code for the head; chordin/Bmp-4 signalling for the dorsoventral organization; *pax6* doing the photoreceptor; *myoD* doing the muscle and so on. It is tempting to imagine this robust ancestor of ours as being one of the creatures that created the fossilized burrows in the Vendian and Cambrian sediments, some of which can still apparently be found complete with fossilized faeces.

The best parts of the book are those that use information from molecular biology to tell us what actually happened in evolution. For example, milk is a novelty created by the mammals. But the lactose synthetase consists of two components: a galactosyl transferase that was formerly used to assemble glycoprotein carbohydrate chains; and α -lactalbumin, which is a modified version of lysozyme. Similar molecular insight is applied to other examples, such as the origin of myelin, of intermediate filaments, and of lens crystallins.

The lens serves to introduce what must be the most amazing animal in the book. The cubomedusoid jellyfish, *Tripedalia*, apparently has image-forming eyes but no central nervous system to process or interpret the image. Creationists used to tell us that the image-forming eye could not possibly have evolved by chance because of all the coordinated changes that would simultaneously have been required in the brain. So I am left unsure whether *Tripedalia* represents real evidence against the existence of God, or just flexible robustness run riot.

The book is not designed for undergraduate teaching as it lacks the basic descriptive embryology on which to hang the molecular story. It is well written but fairly long and the intended market is presumably graduate courses for students already familiar with developmental mechanisms who want to explore their implications for evolution. It is just the job for students who want to make their debating skills in molecular evolution more robustly flexible. □

J. M. W. Slack is in the Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK.

Earthquakes and elephants

Fieldwork: A Geologist's Memoir of the Kalahari

by Christopher Scholz
Princeton University Press: 1997. Pp. 190.
\$24.95, £19.95

Keith Cox

In 1974 Christopher Scholz and his team carried out a survey of seismicity in the Kalahari Desert in southern Africa, at the request of the United Nations Food and Agriculture Organization. They achieved some decent scientific results, but also had a whale of a time, with experiences varying from the comic through the

awe-inspiring to the downright frightening.

Few Earth scientists write anything in the style of their life's memoirs, so this book is doubly welcome. It should appeal to a wide variety of readers, whether fieldworkers or not. The science is accessibly laid out and richly embroidered with tales of the bush.

The scientific problem that the team tackled was to discover whether there is an active extension of the East African rift system into Botswana. Is this the tip of the system propagating itself southward? The question is potentially important because when a fault, such as that forming the edge of a rift, moves and generates an earthquake, there is a change of elevation along the line where the fault-plane reaches the surface. The Kalahari is very flat and the drainage system is in a delicate balance, around the Okavango delta, for example. A large change in the drainage pattern could easily be induced by only a minor movement, and lead to profound ecological consequences.

Botswana is not noted for big earthquakes, but any seismically active area produces many small earthquakes. So the survey had to deal with micro-earthquakes which, predictably, would turn up in sufficient numbers during the few months spent in the field. The technique is to install an array of three or more seismometers with recording devices, leave them for a day or several days, and then see what you have caught. Then the array is moved somewhere else, and so on.

But much of the Kalahari is covered with more-or-less unconsolidated sand, about the worst possible material through which to try to detect micro-earthquakes. As a result, much time had to be devoted to the search for areas of solid bedrock. This travelling about, setting up camp, overcoming obstacles, coping with the wildlife and, not least, confronting officialdom, forms the substance of the book. It is rich in accounts of the incidents that such a mode of life throws up.

It was necessary, for example, to set up camp in thick bush half a mile from the only watering hole for miles. The only clear strip of bush to camp on turned out to be the main route used by elephants at night on their way to have a drink. Add a few thousand nearby antelopes, lions, hyenas and so on, and the night becomes alarmingly noisy.

In the end, the party managed to observe a sufficient number of micro-earthquakes to confirm their hypothesis. The author comments that when the work was published, it did not cause a great stir, but he regards it as an honest and useful job well done.

Although much of the book is devoted to the sheer joy of life in the bush (and its per-

New in paperback

The Lives to Come: The Genetic Revolution and Human Possibilities

by Philip Kitcher
Penguin, £8.99

"Stimulating, informative and courageous.... Kitcher's arguments are always compassionate, always ingenious and always worth pursuing", wrote John Harris in a review in *Nature* 380, 591 (1996).

The Value of Life: Biological Diversity and Human Society

by Stephen R. Kellert
Island Press, \$16.95

"Immensely stimulating, tending to inspire new questions with every one it answers". Thomas Lovejoy, *Nature* 382, 594 (1996).

Our Evolving Universe

by Malcolm S. Longair
Cambridge University Press, £14.95

"A beautifully produced book, in an oversize format, with an elegant typographical design, clear diagrams and many colour illustrations... the text is unflinchingly direct and accurate.... Its content is that of a first-rate introductory course". William Press, *Nature* 382, 220 (1996).

ils), and is written so that you can almost smell the smoke of the camp-fire, the descriptions of occasional trips to town are just as evocative of Africa. We meet a rich array of ramshackle bars with ramshackle customers, we play plenty of darts and hear many a comic or curious yarn.

Perhaps the best is the one about the Afrikaner who, at a time of severe floods, managed one moonless night to drive across a bridge that not only had no hand-rail but was under two feet of water. On being asked how he managed this, he replied: "What bridge?" □

Keith Cox is in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3LY, UK.

Fringe maths

Fermat's Last Theorem: The Story of a Riddle that Confounded the World's Greatest Minds for 358 Years

by Simon Singh

Fourth Estate: 1997. Pp. 362. £12.99

Keith Devlin

On 23 June 1993, the world of mathematics was electrified by the news that the British mathematician Andrew Wiles had proved Fermat's last theorem, a problem in number theory that had resisted numerous attempts at solution for more than 350 years.

Within a few months of that dramatic announcement, an error had been found in the proof. Although everyone agreed that Wiles had produced some of the most remarkable mathematics of this century, he had not, as he had first thought, solved what was almost certainly the most famous open problem in mathematics.

Then, in October 1994, Wiles announced that he had corrected the mistake. This time, the experts agreed that his proof was correct. The complete argument was published in *Annals of Mathematics* in May 1995.

That, in a nutshell, is the story. John Lynch, editor of the BBC 'Horizon' series, sensed a good television documentary and was on to the story from the first announcement. As his co-producer and director, Lynch chose Simon Singh, a physics PhD who had worked in television for five years. The resulting programme was broadcast in 1996. This book is a spin-off from the programme. It comes close on the heels of another book of the same title, by Amir Aczel, which was reviewed in *Nature* last October (383, 774; 1996). That review applies virtually word for word to the Singh book.

Both books adopt the same overall approach. They offer plenty of discussion of the early history of Fermat's last theorem, a selection of mathematics' most durable and

oft-repeated anecdotes and examples, many of them having little connection to the theorem (such as sphere packing, the four-colour theorem, the Galois saga), capped off with the briefest of thumbnail sketches of Wiles's proof.

The similarity is an inevitable consequence of the difficulty facing anyone who sets out to write a popular account of Wiles's solution to the problem of the theorem. The problem itself is easily understood, and the human drama that surrounds it has all the elements of good drama, from the problem's origin right up to the final acknowledgment that it had indeed been solved. But the solution is something else. There are probably only a dozen mathematicians in the world who are sufficiently knowledgeable to follow Wiles's proof fully, and maybe a few hundred who can, with difficulty, struggle through it.

Fermat's claim was that for no whole number n greater than 2 can the equation $x^n + y^n = z^n$ have any whole number solutions for x, y, z (except where one of the numbers is zero).

Wiles's proof involves building a bridge to connect two areas of deep mathematics that had hitherto been regarded as entirely separate, the theory of elliptic functions and the theory of modular forms. As long ago as 1954, two Japanese mathematicians, Yutaka Taniyama and Goro Shimura, had suggested that these two areas were connected. But, although over the years an increasing number of mathematicians had come to accept the possible truth of the Taniyama-Shimura

conjecture, most thought that a proof was many decades away.

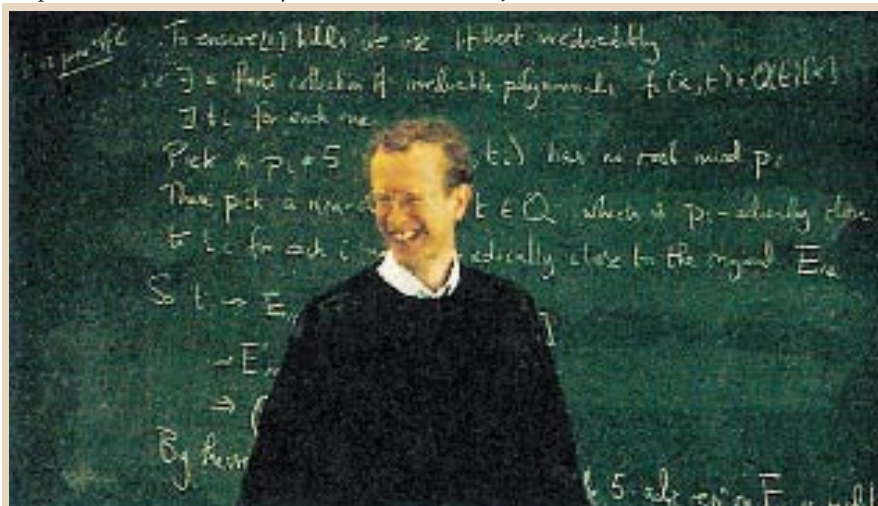
In 1986, Ken Ribet proved the surprising result that Fermat's last theorem is a consequence of the conjecture, providing Wiles with a possible way to attack a problem that had fascinated him since childhood.

By proving a special case of the Taniyama-Shimura conjecture, Wiles not only proved Fermat's last theorem, he opened the doors wide to an entirely new era in number theory. His proof was both an end to a 350-year-old riddle and the beginning of an era in which we will learn much more about those oh-so-familiar counting numbers that lie at the heart of mathematics.

For a would-be author of a book that sets out to explain the story of the last theorem, the only viable approach is to provide mathematical detail around the edges, where its only function is to put some mathematical content between the book's covers, and to concentrate on the human drama when it is time to describe the problem's final resolution. Although I think Singh's book would have been better had he included less of the peripheral 'popular mathematics' dressing, his treatment of the human drama is well done, and as a result the book succeeds.

Non-mathematicians should get something of the flavour of the way mathematicians work, while mathematicians will enjoy learning some of the human details surrounding the eventual solution. □

Keith Devlin is at Saint Mary's College of California, Moraga, California 94575, USA.



Money in the equation

Wiles's proof of Fermat's last theorem is also now described in the paperback edition of Keith Devlin's *Mathematics: The Science of Patterns* (W. H. Freeman/Scientific American Library, \$19.95, £14.95). The book was originally published in 1994, shortly before the proof's discovery, and was reviewed in *Nature* 373, 206 (1995). As Devlin points out, the offer of various awards to the first person to find a

proof has added to the theorem's allure: in 1816, the French Academy offered a gold medal and a cash prize; and in 1908 the Royal Academy of Science in Göttingen offered another cash prize, the Wolfskehl Prize, now worth £30,000. Wiles, pictured here announcing his solution at the Newton Institute in Cambridge on 23 June 1993, receives the latter award tomorrow (27 June).

K⁺ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons

Dax A. Hoffman, Jeffrey C. Magee*, Costa M. Colbert & Daniel Johnston

Division of Neuroscience, Baylor College of Medicine, Houston, Texas 77030, USA

Pyramidal neurons receive tens of thousands of synaptic inputs on their dendrites. The dendrites dynamically alter the strengths of these synapses and coordinate them to produce an output in ways that are not well understood.

Surprisingly, there turns out to be a very high density of transient A-type potassium ion channels in dendrites of hippocampal CA1 pyramidal neurons. These channels prevent initiation of an action potential in the dendrites, limit the back-propagation of action potentials into the dendrites, and reduce excitatory synaptic events. The channels act to prevent large, rapid dendritic depolarizations, thereby regulating orthograde and retrograde propagation of dendritic potentials.

Neuronal dendrites contain a variety of voltage-gated Na⁺ and Ca²⁺ channels that enable action potentials and excitatory postsynaptic potentials (EPSPs) to propagate actively throughout the neuron (reviewed in refs 1, 2). Although our understanding of information processing in neurons is increasing, several fundamental aspects of dendritic physiology remain unexplained. Until now, reports of dendritic action-potential threshold and amplitude, as well as the small, active EPSP component, have suggested that the dendritic membrane is weakly excitable compared with the somatic membrane^{3–11}, which is inconsistent with the comparable density of Na⁺ and Ca²⁺ channels in these two regions^{3–5}.

Our elementary knowledge about dendritic K⁺ channels^{10,12} means that our understanding of dendritic electrical properties is incomplete. We show here that the dendrites of CA1 pyramidal neurons have a high density of transient A-type K⁺ channels and that this density increases with distance from the soma. The presence of these dendritic channels causes action-potential amplitude to decrease with distance from the soma, inhibits the ability of dendrites to initiate action potentials, and alters the shape of EPSPs. Furthermore, several key voltage-dependent properties of these channels allow for the unique regulation of dendritic excitability by subthreshold synaptic activity. Such a feature could be important for the induction of associative synaptic plasticity^{13–15}. By severely dampening dendritic excitability, A-type K⁺ channels play the dominant role in determining the electrical properties of CA1 dendrites and thereby influence dendritic integration and propagation of information.

Dendritic density of transient K⁺ channels

To characterize the properties and distribution of K⁺ channels in a central neuron, we have used cell-attached patch-clamp recordings from the soma and apical dendrites of CA1 hippocampal pyramidal neurons. These recordings revealed a high density of outward current composed of two distinct components separated through the use of standard subtraction procedures¹⁶. The first component was a transient component that rapidly activated and inactivated; the second component was more slowly inactivating. The recorded density of the transient component increased linearly with distance from the soma (slope, 12.5 pA per 100 μm) (Fig. 1a). The peak

current density of transient channels increased from an average of 8.32 ± 1.38 pA per patch (mean $\pm$ s.e.m.) in the soma to 51.56 ± 9.47 pA in the distal dendrites for a voltage step from -85 to 55 mV. The current density of the sustained, non-inactivating component remained constant, with an average of 8.39 ± 1.46 pA per patch in the soma and 9.55 ± 1.84 pA per patch in the distal dendrites. Representative traces from the soma and distal dendrite of both the transient and sustained component are shown in Fig. 1b. Single channel recordings such as those in Fig. 1c were used to construct the *I*–*V* curve in Fig. 1d. A linear-regression fit of the data reveals that the transient component comprises small conductance (~ 7.5 pS) channels whose unitary current amplitude reverses near the K⁺ equilibrium potential.

Steady-state activation and inactivation curves were generated for both the transient and sustained components (Fig. 2a,b). These curves reveal that the transient component activates at slightly more hyperpolarized potentials than the sustained component, and that significant transient-channel activation will occur at relatively hyperpolarized potentials (compare Fig. 2a,b). This particularly applies to transient channels located in the distal dendrites, which had an activation curve that was shifted 12 mV hyperpolarized compared to that from the soma/proximal (up to 100 μm) dendrites (Fig. 2a). The time course of activation was rapid (~ 1 ms) and was relatively constant over the range of voltages tested. The steady-state inactivation curves for the transient component did not vary with distance from the soma (Fig. 2a), and the sustained component had a steady-state activation curve that was also independent of patch location (Fig. 2b). The time constant of inactivation for the transient component increased linearly with amount of depolarization from 6.1 ± 0.4 ms for a step to -25 mV up to 27.3 ± 1.9 ms for a step to $+55$ mV (Fig. 2c). Similar results were obtained using outside-out patches ($n = 3$, data not shown). Thus, both an increased transient current density and a more hyperpolarized activation range should increase the impact of these channels on more distal dendrites.

4-Aminopyridine (4-AP), tetraethylammonium (TEA), and α -dendrotoxin (DTX) were applied to outside-out patches pulled from the soma and dendrites. In no case did the location from which the patch was taken have a discernible effect on drug efficacy. 4-AP reduced the transient component in a dose-dependent manner while having only a slight antagonist effect on the sustained component (Fig. 2d). TEA reduced both components similarly at

* Present address: Neuroscience Center, Louisiana State University Medical Center, New Orleans, Louisiana 70112, USA.

0.1 and 1.0 mM but had a larger effect on the sustained component at 10 mM (Fig. 2e). DTX has been reported to preferentially reduce a transient K^+ conductance with kinetics and 4-AP sensitivities separate from the transient A-type channels^{17–19}. DTX (1 μ M) reduced the transient component by $21.83 \pm 6.67\%$ ($n = 6$; data not shown).

The voltage-activated outward current described above appears to be composed of at least three components. The primary component is a large, transient, DTX-insensitive current that possesses voltage-dependent and pharmacological properties most similar to those generally ascribed to A-type K^+ channels^{16–22}. A small fraction of the transient current was DTX-sensitive, which may indicate a second channel type. Finally, the sustained component demonstrated voltage-dependent and pharmacological properties similar to delayed-rectifier-type K^+ channels. The elevated density of dendritic A-type channels reported here agrees with immunohistochemistry studies that found an increased density of the transient K^+ channel subtype Kv 4.2 in the distal dendrites of CA1 pyramidal neurons^{23,24}. Another recent study has found Kv 4.2 clustered on the postsynaptic membrane directly apposed to the presynaptic terminal²⁵.

Role of A-channels

A large outward conductance that both activates and inactivates very rapidly over relatively hyperpolarized voltage ranges could potentially have a profound impact on the electrical properties of dendrites. Such a conductance would act like a shock absorber to limit large, rapid dendritic depolarizations. CA1 and neocortical pyramidal cell types contain voltage-gated Na^+ and Ca^{2+} channel densities that are uniform for what appears to be the entire extent of the somatodendritic axis^{3,5}. The increasing density of dendritic A-channels could explain why, in the face of this uniform inward conductance, the amplitude of back-propagating action potentials

decreases with distance from the soma, why action potentials initiate in the axonal–somatic region as opposed to the dendrites, and why dendritic Na^+ and Ca^{2+} channels seem to have only a slight impact on subthreshold EPSPs. To assess the functional impact of the increasing density of A-type channels in the dendrites, simultaneous dendritic and somatic whole-cell voltage recordings were made from CA1 pyramidal neurons in hippocampal slices. Blockade of mainly A-type K^+ -channels, by bath application of 4-AP (3–10 mM; Fig. 2d), increased dendritic action-potential amplitude in a concentration-dependent manner (Fig. 3a,e)¹⁰. Under such conditions, very little amplitude attenuation was observed as action potentials propagated into the dendrites. Single action potentials were transformed into bursts of 2–3 spikes or plateau potentials that led to nearly constant depolarization in the dendrites and continuous repetitive firing in the soma (Fig. 3a,b). Addition of 200 μ M $CdCl_2$ or low- Ca^{2+} external solution (0.5 mM Ca^{2+} , 8 mM Mg^{2+}) blocked the generation of plateau potentials and burst firing, indicating that dendritic Ca^{2+} channel activation is required for such activity (Fig. 3a). Upon blockade of transient dendritic K^+ channels, action-potential duration, time to peak, and amplitude were increased, whereas peak rate of rise was not significantly changed (Fig. 3e). These data indicate that blocking dendritic A-type channels may allow dendritic action potentials to continue to depolarize until they reach a peak voltage similar to that of the somatic action-potential. These large-amplitude, back-propagating action potentials activated dendritic Ca^{2+} channels, resulting in a massive influx of Ca^{2+} into the dendrites (Fig. 3a,c). Together our results provide strong evidence that an increased density of A-type K^+ channels in the dendrites of CA1 pyramidal neurons acts to reduce action-potential amplitude and limit the back-propagation of full-amplitude action potentials to only the most proximal dendritic regions. Furthermore, the elevated density of dendritic A-channels does not allow single somatic action potentials to induce

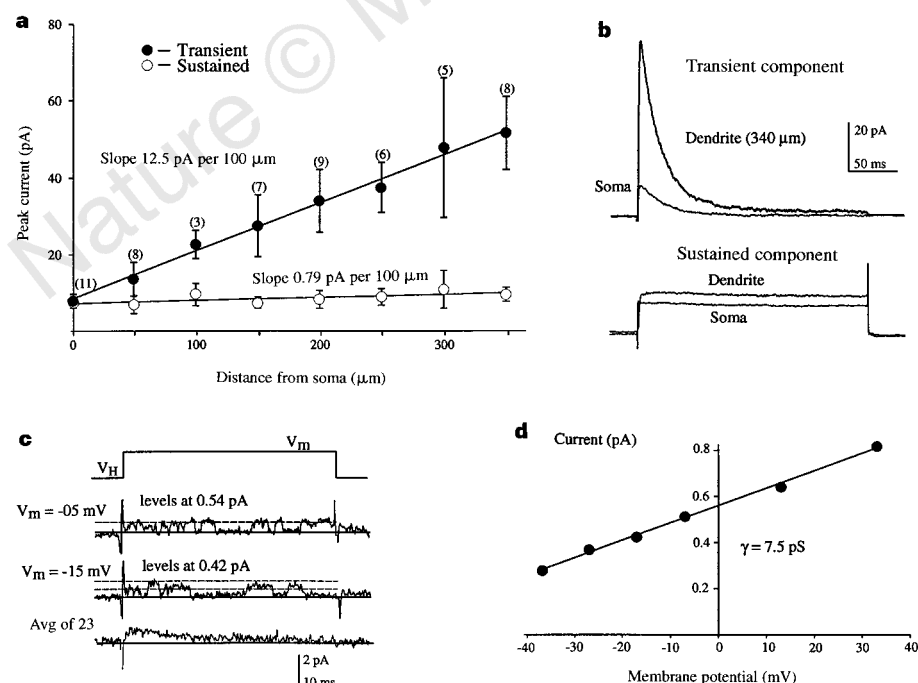


Figure 1 High density of transient K^+ channels in distal dendrites. **a**, The peak outward current amplitude for the transient and sustained components in response to a voltage step from -85 to $+55$ mV in 11 somatic and 46 dendrite-attached patches (dendrite patches are binned to 50- μ m segments). Best fit lines are shown for both components. **b**, Representative traces for both components from the soma and from the dendrites 340 μ m from the soma. The two components were separated out from the total outward current by a standard subtraction procedure¹⁴. **c**, Single-channel recordings of transient A-type channels from cell-attached patches. The patches were voltage-clamped at holding potentials of -65 and -55 mV and stepped to -15 and -5 mV, respectively. Unitary amplitudes are shown. The bottom trace is the ensemble average of 23 sweeps to -15 mV. **d**, I/V plot of single transient K^+ channels. Unitary current amplitudes are plotted as a function of membrane potential from four separate experiments (error bars are smaller than the symbols). The mean slope conductance is 7.5 pS.

bursts or repetitive firing, thus maintaining the linear relationship between action-potential firing rate and Ca^{2+} influx^{26–29}.

The blockade of dendritic A-type channels also allowed action potentials to be initiated in the dendritic compartment (Fig. 3d). In the presence of 4-AP, dendritic current injections or large-amplitude EPSPs could initiate action potentials locally in the dendrites. The effect was likely to be the result of a decrease in dendritic threshold and an increase in the amount of depolarization induced by a given current injection. Although the dendritic membrane was capable of initiating action potentials when A-channel density was reduced, the somatic–axonal region still appeared to have the lowest threshold, because small long-duration dendritic current injections (0.3 nA, 300 ms) evoked axonally initiated action potentials. Thus, A-channel blockade increased the excitability of the dendritic membrane, allowing for the initiation and propagation of fully regenerative action potentials in the dendrites.

The high density of dendritic A-type channels had substantial effects on subthreshold synaptic events propagating through the dendrites. Bath application of 4-AP increased the amplitude, time to peak, and duration of EPSP-shaped voltage transients induced by dendritic current injections (Fig. 4a,b,e). This effect was similar for both somatic and dendritic compartments. Upon wash-in of 4-AP, subthreshold dendritic current injections (~7 mV in soma) were transformed into suprathreshold (>11 mV in soma) bursts of action potentials. In some cases, the action potentials were initiated in the dendritic compartment. As with action-potential shaping, A-channel blockade allowed EPSPs to depolarize further than would occur with a full complement of A-channels present^{12,30}. Application of tetrodotoxin (TTX, 300–500 nM) after A-channel blockade revealed that a large fraction of the 4-AP-induced increase in depolarization was due to Na^+ channel activation (Fig. 4c). Thus, A-channels act to counter additional inward current that is induced by subthreshold Na^+ channel activation, effectively eliminating

any Na^+ channel EPSP boosting. Although dendritic A-channels reduce the amount of depolarization provided to the soma by synaptic input, the relative amount of voltage decrement occurring from the dendrite to the soma was unchanged (control decrement equalled $52 \pm 5\%$ compared with $55 \pm 5\%$ with 4-AP) (Figs 4d,e and 5a).

Computer simulations

The whole-cell recordings demonstrate that an increasing density of dendritic A-channels determines action-potential amplitude and threshold, shapes EPSP propagation through the dendrites, and decreases overall membrane excitability. These results show that dendritic Na^+ channel density is sufficient to generate (and even initiate) full-amplitude action potentials once the high A-channel density is reduced. We have used computer simulations to illustrate how this complement of dendritic voltage-gated channels might interact to produce the observed whole-cell properties. Upon constructing a reduced model that contained Na^+ and K^+ channels with the characterized properties, we considered the effect of dendritic voltage-gated channels on the spread of EPSPs towards the soma. The model confirmed that the high density of A-channels may counteract EPSP boosting provided by subthreshold Na^+ -channel activation, so that EPSP amplitude is smaller than if the membrane were passive (Fig. 5a). Thus the net effect of the dendritic voltage-gated channels at the estimated densities is to decrease the depolarization provided by the synaptic input. Furthermore, the largest proportion of channel activation, and therefore active current flow, is in the dendritic regions where the synaptic input is located. A proportionally smaller amount of channel activation occurs as the EPSP moves into the more proximal and less depolarized regions, so that they eventually are propagating almost passively. These observations suggest that most EPSP shaping occurs locally at the site of input and helps explain why

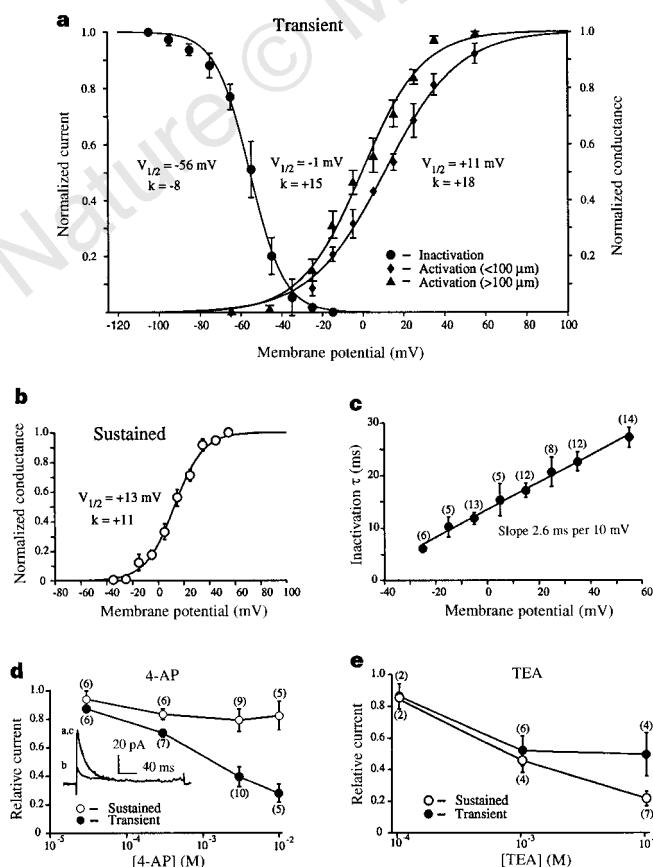


Figure 2 K^+ channel voltage-dependent properties and pharmacology. **a**, Activation and inactivation curves of transient K^+ channels recorded from soma- and dendrite-attached patches. The activation curve for distal dendrites (>100 μm) (triangles) is shifted hyperpolarized compared to that of somatic/proximal dendrites (diamonds) ($V_{1/2} = -1$ and $+11$ mV, respectively), although their slopes are similar (k is 15 and 18, respectively). The inactivation curve had a $V_{1/2}$ of -56 mV and a k value of -8 (n is 3 to 12). **b**, Activation curve of the sustained component ($V_{1/2} = 13$ mV, $k = 11$, $n = 4$ to 9). **c**, The time constant of inactivation (τ) for the transient channels increases with depolarization. τ is plotted against command potential and fit by a regression line with a slope of 2.6 ms per 10 mV. **d**, Effect of 4-AP on the transient (filled circle) and sustained (open circle) components of the total outward current upon a voltage step from -80 to $+40$ mV in outside-out patches pulled from the soma and dendrites. 4-AP blocked 50% of the transient current at about 1.4 mM. At no concentration did 4-AP affect more than 20% of the sustained component. Inset: a, trace showing total outward current before application of 4-AP; b, after application of 10 mM 4-AP; c, after washout of 4-AP. **e**, Effect of TEA on the transient (filled circles) and sustained (open circles) components of the total outward current on a voltage step from -80 to $+40$ mV. TEA at 1 mM blocked about 56% of the sustained current and 50% of the transient current. At 10 mM TEA blocked 80% of the sustained but still only about 50% of the transient current.

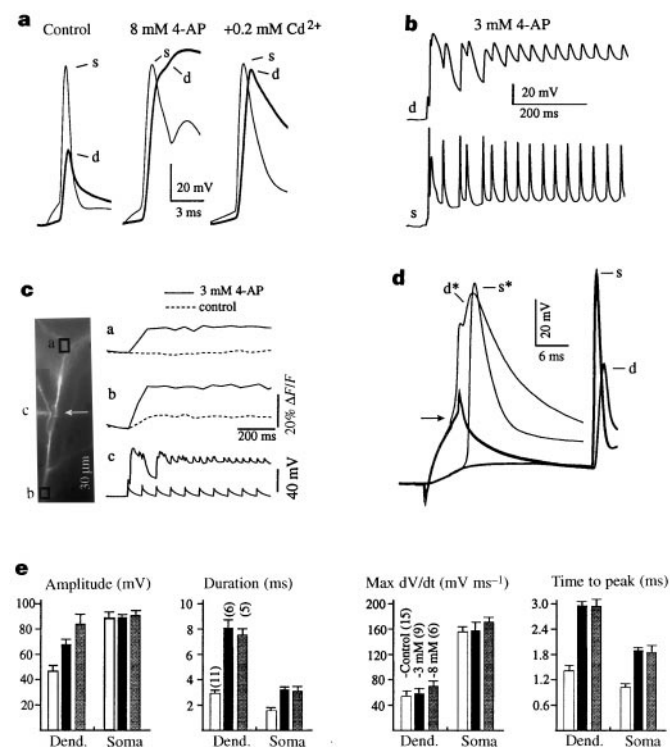


Figure 3 Effect of A-channel blockade on dendritic action potentials. **a**, Under control conditions, current injection into the soma (300 pA, 10 ms) evoked an action potential (somatic recording labelled s) that was severely attenuated upon reaching the dendritic recording pipette ($\sim 250 \mu\text{m}$, labelled d). Bath application of 8 mM 4-AP caused dendritic action potential amplitude to increase dramatically and a slower plateau potential was also generated. Subsequent addition of $200 \mu\text{M}$ CdCl_2 reduced the duration of the dendritic action potential without affecting the amplitude. **b**, In the presence of normal external Ca^{2+} and 3 mM 4-AP, a single dendritic current injection (1.3 nA, 2 ms) led to prolonged depolarization in the dendrites and repetitive spiking in the soma. Note that the somatic compartment was better able to repolarize than the dendritic compartment. **c**, Fura-2 Ca^{2+} imaging of distal dendrite during a 20 Hz train of back-propagating action potentials. Application of 3 mM 4-AP transformed a moderate and spatially graded influx of Ca^{2+} influx (control; dashed lines) into a large global Ca^{2+} influx (4-AP; solid lines). Traces to the right of the image correspond to the average percentage change in fluorescence ($\Delta F/F$) from the regions inside the appropriately labelled boxes. Bottom trace is dendritic (white arrow) voltage recording for control (smaller action potential) and in 4-AP (plateau) potential. Note that the fidelity of action-potential back-propagation is severely reduced by the generation of a Ca^{2+} -dependent plateau potential during A-channel blockade. **d**, Under control conditions (heavy traces) a 6 ms, 0.9 nA current injection into the dendrite evoked an axonally initiated action potential with a ~ 15 ms delay. Following 4-AP (8 mM, light traces) wash-in, the same current injection evoked an action potential that was initiated in the dendrite. Arrow points out that apparent threshold was lowered by A-channel blockade. **e**, Plots of action-potential amplitude, duration (at half maximum voltage), maximum rate of rise, and time to peak for control (white bars), 3 mM 4-AP (black bars), and 8 mM 4-AP (grey bars). Number of cells in each group is shown in the maximum rate of rise plot, except for those used for the duration measurements, which are shown separately. The duration measurements include only those neurons that were recorded under conditions of low Ca^{2+} influx (that is, in Cd^{2+} or low Ca^{2+} solutions).

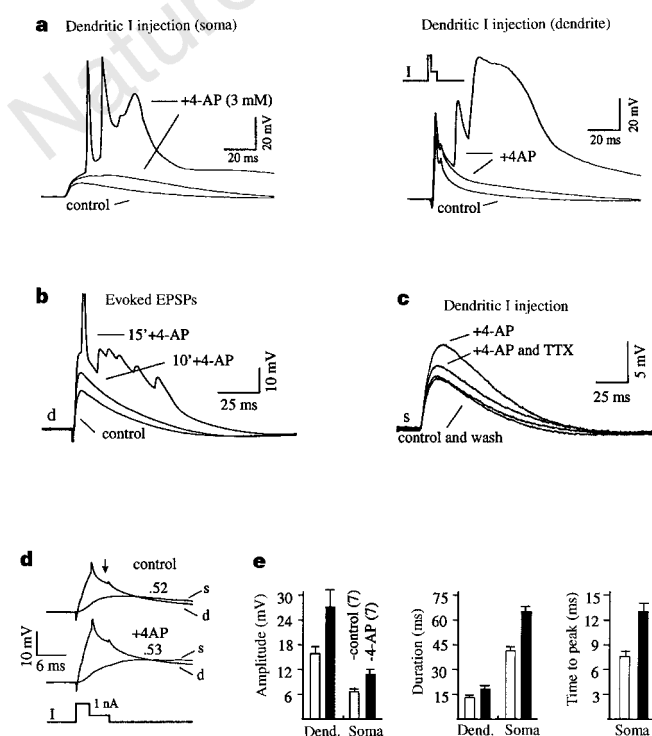


Figure 4 Effect of A-channel blockade on the propagation of EPSPs. **a**, EPSP-shaped voltage transients were produced by a two-step current injection into the dendrite, which under control conditions was subthreshold (~ 7 mV in the soma). Wash-in of 3 mM 4-AP caused a dramatic increase in the duration and amplitude of the voltage transients in both the somatic (left set of traces) and dendritic (right set of traces) compartments. Upon complete wash-in, the previously subthreshold current injection evoked a burst of action potentials. **b**, Synaptically evoked EPSPs during control conditions, 10 min and 15 min after perfusion of the dendritic electrode with an internal solution containing 6 mM 4-AP. Previously subthreshold EPSP became suprathreshold evoking bursts of action potentials. Each trace is an average of 5 EPSPs evoked by Schaffer collateral stimulation at 0.25 Hz. **c**, Subthreshold Na^+ channel activation contributes to EPSP amplitude. Bath application of 300 nM TTX subsequent to the application of 3 mM 4-AP, substantially reduced the amount of EPSP amplification induced by A-channel blockade. **d**, The relative amount of EPSP attenuation from dendrite to soma was unchanged by 4-AP application ($\sim 50\%$ in both cases). Traces are superimposed dendritic and somatic voltage transients induced by dendritic current injections under control and 4-AP conditions. Note that the amplitude of the dendritic and somatic EPSP increased proportionally with A-channel blockade. Arrow indicates where amplitude measurement was taken from dendritic transient. **e**, Plots of EPSP waveform amplitude, duration (at half maximum voltage), and time to peak for control (white bars) and 3 mM 4-AP (black bars). Number of cells in each group is shown in the amplitude plot.

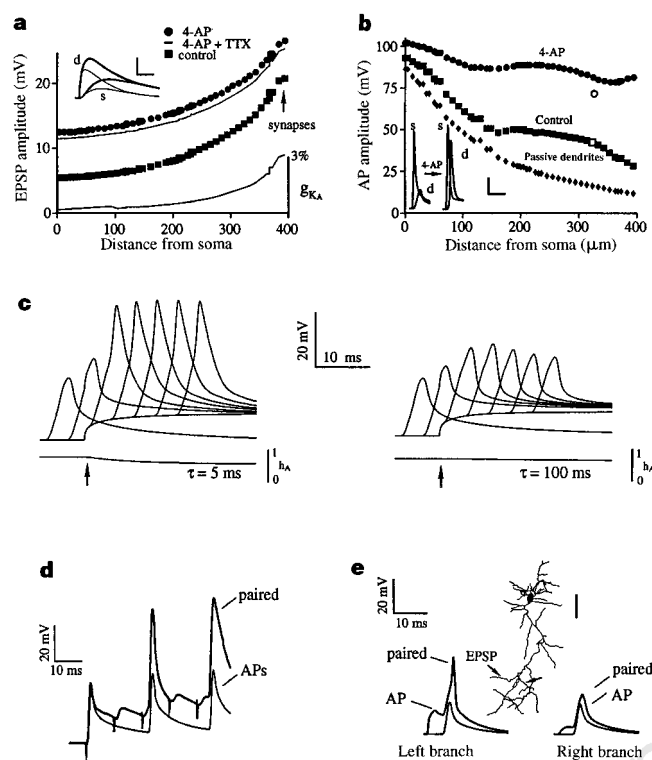


Figure 5 Computer simulations. A-channels shape dendritic EPSPs and action potentials. **a**, A-channels decrease EPSP amplitude primarily at site of input. Peak depolarization of dendrites along path from synaptic input at 390 μm to soma using reconstructed neuron in **e**. Upper plot: normal densities of Na^+ and A-channels (squares), 10% A-channel density (circles) to simulate 4-AP application, and 10% A-channels and no Na^+ channels (solid line) to simulate 4-AP + TTX. Lower plot: peak conductance of A-channels with distance during the EPSP. Note that the conductance decreased rapidly with distance. Inset: waveforms are EPSPs in dendrites (larger pair) and soma (smaller pair) with normal densities (light) and with 10% A-channel density (bold). Scale: 10 mV, 5 ms. **b**, A-channels determine action-potential amplitude. Dendritic action potential amplitude versus distance from soma: normal A-channel density (squares, control), 10% A-channel density (circles, '4-AP'), 10% A-channel density and no dendritic Na^+ channels (diamonds, passive dendrites). Left inset: pairs of somatic (large) and dendritic (small) action-potential waveforms with full A-channel density (left) and 10% A-channel density (right). Decreased A-channel density restricted to a single dendritic branch (open circle) locally increased the amplitude compared to control (open square). Scale: 20 mV, 5 ms. **c**, Inactivation of A-channels increases back-propagating dendritic action-potential amplitude. To inactivate A-channels the dendrite was depolarized (beginning at the arrow) from -67 mV to -55 mV, which decreased the available channel population (h_A) from ~ 0.8 to ~ 0.5 with $\tau_h = 5$ ms. Back-propagating action-potential amplitude increased with increasing duration of the preceding depolarization. With the τ_h of inactivation slowed to 100 ms for potentials below -53 mV, (h_A) was decreased only slightly by the depolarization. With increasing duration of preceding depolarization, dendritic action-potential amplitude did not continue to increase. **d**, EPSPs paired with back-propagating action potentials increase dendritic action-potential amplitude. Actual whole-cell recording (not a simulation) from ~ 240 μm from soma. Action potentials were evoked by 2-ms current injections through a somatic whole-cell electrode at 20-ms intervals. Alone, action-potential amplitude was small (APs). Paired with EPSPs, the action-potential amplitude increased greatly (paired). **e**, Local depolarization selectively increases dendritic action-potential amplitude. Reconstructed pyramidal neurons with symmetrical left and right branches (arrows). Alone, the back-propagating action potential was small and of similar amplitude in both branches (AP). With synaptic input to the left branch (EPSP), the amplitude of the back-propagating action potential increased greatly in the left branch (paired) but not in the right branch (paired). Scale bar, 100 μm .

the relative amount of voltage decrement occurring from the dendrite to the soma is similar with or without dendritic voltage-gated channels.

We next turned to the propagation of action potentials into the dendrites. Previously, the decline of dendritic action-potential amplitude has been attributed to the presence of a density of Na^+ channels that, in everywhere but the axon, is incapable of sustaining a regenerative action potential^{31,32}. As shown in Fig. 3, however, A-channel blockade allowed dendritic action potentials to become fully regenerative. In our simulations, incorporation of dendritic A-channels allowed the neuron to produce large-amplitude somatic action potentials that declined as they propagated into the dendrites (Fig. 5b, squares). Decreasing the A-channel density to 10% of its normal value resulted in the action-potential amplitude remaining large throughout the dendritic arbor. Thus, an increasing density of dendritic A-channels determines final action-potential amplitude even when the dendrites contain a uniform Na^+ channel density that can generate fully regenerative action potentials.

We also noted that decreasing the A-channel density in only a single dendritic branch of the model dramatically increased the amplitude of the action potential only in that dendritic compartment (open circle in Fig. 5b). This simulation shows that local modulation of A-channel density can affect membrane excitability in a limited region of the neuron. A-type, K^+ -channel activity is controlled by numerous neurotransmitters, chemical messengers, cation concentrations, auxiliary subunits, and by oxidative and phosphorylation states^{33–39}. Localized application of such modulatory substances, for example during transmitter release, could lead to a highly local alteration in dendritic membrane excitability. The vast number and diversity of agents that affect these channels may be indicative of their importance in regulating membrane excitability.

A-channel density can also be regulated through membrane depolarization. The voltage-dependent inactivation properties of A-type K^+ channels (Fig. 2) indicate that modest levels of membrane depolarization could decrease the size of the available A-channel population and likewise increase dendritic action-potential amplitude. Such a mechanism would allow subthreshold synaptic activity to regulate dendritic action-potential propagation^{13,40}. We used the model to test the hypothesis that moderate dendritic depolarization would significantly inactivate A-type K^+ channels (Fig. 5c) and thus allow dendritic action-potential amplitude to increase. A dendritic compartment was depolarized (arrow in Fig. 5c), and back-propagating action potentials were initiated in the axon before and at progressively later times during the depolarization. During the depolarizations (~ 12 mV), the fraction of available A-channels (h_A) decreased, and action-potential amplitude increased accordingly. To control for other variables that might influence the amplitude of the action potential, we modified the A-channel model so that inactivation proceeded very slowly ($\tau = 100$ ms) at the steady-state dendritic depolarization, but would inactivate normally when repolarizing the action potential. Thus, during the depolarizing step, the available channels remained essentially constant until the action potential invaded. With these modified A-channels, the amplitude of the action potential increased somewhat with the depolarization (reflecting the increase in Na^+ channel activation), but did not continue to increase throughout the depolarization. Action-potential amplitude decreased later in the depolarization as the Na^+ channels inactivated. Thus, rapid, depolarization-induced inactivation of A-channels is a plausible mechanism for increasing action-potential amplitude in the dendrites.

Under physiological conditions such moderate membrane depolarization can be provided by subthreshold synaptic input. In fact, when back-propagating action potentials are paired with subthreshold synaptic depolarizations, action-potential amplitude (as well as evoked Ca^{2+} influx) is increased in a supralinear fashion^{13,41}.

(Fig. 5d). One possible explanation for this amplitude boosting is that the depolarization provided by EPSPs inactivates A-channels and decreases their effect on the amplitude of the action potentials. Furthermore, if the local differences in membrane depolarization that occur during synaptic input could similarly affect dendritic excitability, synaptic depolarization might provide a spatially localized amplification of back-propagating action potentials. We simulated this situation by pairing synaptic input onto a single branch of a distal dendrite with back-propagating action potentials. A relatively symmetrical branch point was chosen so that the unpaired action potentials were similar in both daughter branches. When back-propagating action potentials were paired with the localized synaptic input, action-potential amplitude was significantly boosted in the branch receiving synaptic input. As the other branch was only slightly depolarized by the spread of the EPSP, boosting of the action-potential amplitude did not occur. Therefore the susceptibility of A-channels to rapid inactivation during sub-threshold synaptic input could provide a mechanism whereby back-propagating action potentials will preferentially invade synaptically active regions of the dendritic arborization. This synaptically regulated propagation of action potentials into specific dendritic branches may play a role in certain forms of hebbian synaptic plasticity¹³.

Discussion

The unexpected finding that dendrites of CA1 pyramidal neurons contain a high density of transient, A-type K⁺ channels may answer several questions about dendritic function. The presence of these channels could explain why action potentials do not initiate in the dendrites, even during synaptic input, why the amplitude of back-propagating action potentials decreases with distance from the soma, and why there is not a larger boosting of EPSP amplitude, even though inward dendritic conductances are activated. Our results suggest that the linear increase in A-channel density with distance into the dendrites acts, at least transiently, to counter dendritic depolarization. This dampening effect reduces the ability of dendrites to initiate action potentials, decreases the amplitude of back-propagating action potentials, and reduces the magnitude of EPSPs.

Back-propagating action potentials are critical for the induction of associative LTP in CA1 neurons because they allow the output region of the neuron (axon) to communicate with the dendrites^{13–15}. Dendritic Na⁺ and Ca²⁺ channels provide a mechanism for action potentials to back-propagate into the dendrites. If left unchecked, however, inward conductances in the dendrite result in excessive membrane excitability and unrestricted Ca²⁺ influx (Fig. 3b,c), both of which are detrimental to neuron viability and function. The biophysical characteristics of dendritic A-type K⁺ channels make them ideally suited for the task of reducing dendritic excitability. The rapid channel inactivation that occurs at potentials near the resting potential may allow synaptic activity to release local regions of the dendrite from the dampening effect of a high A-channel density. This allows action potentials occurring at the same time as EPSPs to increase in amplitude in a spatially restricted region of the neuron. The inactivation of A-channels by EPSPs and the subsequent amplification of the action potential and evoked Ca²⁺ influx provides a plausible biophysical explanation for the hebbian associativity that is observed in the more distal dendritic regions of these neurons¹³.

Finally, the elevated density of dendritic A-channels actually heightens the associativity provided by back-propagating action potentials. Because A-channel density can be modulated by localized synaptic depolarization, full-amplitude action potentials capable of evoking significant Ca²⁺ influx into the more distal dendrites are only generated in synaptically active regions of the dendritic arborization. Alternative models, such as one in which there is a decreasing Na⁺ channel density in the dendrites, would not have this feature. Thus, dendritic A-channels provide a

dampening effect on dendritic excitability which in turn strengthens the associative link between the input and output regions of the neuron. By shaping action potentials and EPSPs as well as evoked Ca²⁺ influx, A-channels may regulate communication between the soma and dendrites, thereby determining how information is integrated and stored in pyramidal neurons. □

Methods

Hippocampal slices (400 μm) were prepared from 5–8-week-old Sprague–Dawley rats and visualized with a Zeiss Axioskop using infrared video microscopy and differential interference contrast (DIC) optics^{5,41}. All neurons had a resting membrane potential between –57 and –72 mV. For all channel recordings, bath solutions contained (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH₂PO₄, 25 NaHCO₃, 2.0 CaCl₂, 1.0 MgCl₂, 25 dextrose and for outside-out patches 1.0 μM TTX (Sigma). The external solution was bubbled with 95% O₂/5% CO₂ at ~22 °C (pH ~7.4) for all recordings. For cell-attached recordings, the pipette solution consisted of (mM): 125 NaCl, 10 HEPES, 2.0 CaCl₂, 1.0 MgCl₂, 2.5 KCl and 1.0 μM TTX (pH 7.4 with NaOH). For outside-out patch recordings the pipette solution consisted of (mM): 20 KCl, 10 HEPES, 10 EGTA, 120 potassium gluconate, 4.0 Mg₂ATP, 0.3 Tris₂GTP, 14 phosphocreatine and 4 NaCl (pH 7.25 with KOH). Pipettes were pulled from borosilicate glass (10–20 MΩ) and coated with sylgard. Tips were visually inspected before use and had uniform diameters of ~1 μm for both dendritic and somatic recordings. Channel recordings, using an Axopatch 1D amplifier (Axon Instruments), were analogue-filtered at 2 kHz and digitally filtered at 1 kHz off-line. Leakage and capacitive currents were digitally subtracted by averaging null traces or scaling traces of smaller amplitude.

For activation plots, chord conductance, calculated from the peak ensemble current amplitude from a holding potential of –85 mV, was normalized to maximum and plotted as a function of test potential. For inactivation, peak ensemble current amplitude for a step to +55 mV was normalized to maximum and plotted as a function of holding potential. Curves are least-square fits of the data to Boltzmann functions. Inactivation time constants were fit by a non-linear, least-squares program and were well fitted by a single exponential. 4-AP, TEA (Sigma) and DTX (Alomone Labs) were added to a solution identical to the bath solution (except for the 10 mM concentrations, in which case dextrose was decreased from 25 to 15 mM) and was applied to the excised patch by a small-bore perfusion pipette. The sustained component often ran down in outside-out patches, therefore only those experiments where a drug washout was obtained were used to construct the dose–response graphs. Error bars represent s.e.m. and the number of patches (*n*) is given in parentheses.

Whole-cell patch-clamp recordings were made using an Axoclamp 2A (Axon Instruments) and Dagan IX2-700 amplifier in ‘bridge’ modes. The external recording solution contained (mM): 124 NaCl, 2.5 KCl, 1.2 NaH₂PO₄, 25 NaHCO₃, 2.5 CaCl₂, 1.5 MgCl₂, and 10 dextrose, 0.01 DNQX, bubbled as described at ~35 °C (pH ~7.4). Whole-cell recording pipettes (somatic, 2 to 4 MΩ; dendritic, 7 to 10 MΩ) were pulled from borosilicate glass. The internal solution was the same as above, except that EGTA was omitted. Series resistance for somatic recordings was 6–20 MΩ, whereas that for dendritic recordings was 15–50 MΩ. Dendritic pipettes were coated with sylgard. To measure changes in [Ca²⁺]_i, the fluorescent indicator Fura-2 (20–80 μM) was included in the pipette solution, and a cooled CCD camera (Photometrics) used to record changes in fluorescence⁴². When synaptic stimulation was used, methods were similar to those described^{4,42}; DNQX was omitted from the external solution and 10 μM bicuculline was added. Voltages were not corrected for the junction potential (–7 mV) in any of the whole-cell experiments.

Computer simulations were made using NEURON⁴³ with an integration time-step of 0.025 ms. A biocytin-filled hippocampal pyramidal neuron from an adult male rat was reconstructed using NTS (Eutectics; Fig. 5e). Passive electrical parameters were $R_m = 30,000 \Omega \text{ cm}^2$, $C_m = 1 \mu\text{F cm}^{-2}$ in the somatic compartment and $1.6 \mu\text{F cm}^{-2}$ in the dendritic compartments to account for spines. The axial resistivity R_i was $150 \Omega \text{ cm}$, except in the axon where R_i was $100 \Omega \text{ cm}$.

Synaptic input was modelled as a conductance with dual exponential time course of the form $(1 - e^{-t/\tau_1})e^{-t/\tau_2}$, where $t = 0$ at the onset of the conductance ($\tau_1 = 1.5 \text{ ms}$ and τ_2 was 5–10 ms). Synapses had a reversal potential of 0 mV. EPSPs of 15–20 mV near the site of input were produced by 5–7

simultaneously activated synaptic conductances distributed along $\sim 50 \mu\text{m}$ of dendrite. Antidromic spikes were evoked by a brief current injection into the most distal axonal segment.

The model included four voltage-gated channels: Na^+ , proximal and distal A-type K^+ , and DR type. Conductances were set to reproduce the experimentally observed A-type to Na^+ current ratio of $\sim 3:1$ in cell-attached patches from distal dendrites. We compared the magnitude of the A-type current evoked by a step from -90 to $+45 \text{ mV}$ to the magnitude of the Na^+ current evoked by a step from -90 to -10 mV to determine this ratio. The DR-type conductance was set to repolarize the action potential in the absence of any A-type current. Conductances in the model were uniformly scaled up by a factor of 2 to yield a somatic action potential with a peak rate of rise of $\sim 200 \text{ mV ms}^{-1}$. The resulting peak conductances were $\bar{g}_{\text{Na}} = 0.012 \text{ S cm}^{-2}$, $\bar{g}_{\text{KA}} = 0.007 + 0.011$ (distance from soma in $\mu\text{m}/100$) S cm^{-2} and $\bar{g}_{\text{KDR}} = 0.0075 \text{ S cm}^{-2}$. The proximal A-channel model was used within $100 \mu\text{m}$ from the soma, otherwise the distal model was used. $\bar{g}_{\text{Na}}$ and $\bar{g}_{\text{KDR}}$ were uniform in the soma and dendrites and increased 10-fold in the axon. Reversal potentials were $E_{\text{Na}} = +58 \text{ mV}$ and $E_{\text{K}} = -80 \text{ mV}$.

Voltage-dependent time constants (τ) and steady-state (ss) values were used to define the instantaneous values of activation (m) and inactivation (h) gates according to: $dx = (x_{\text{ss}} - x)(1 - e^{-dt/\tau_x})$, where x is a gate variable, and dx is the change in x made during a time step dt .

The Na^+ channel model was of the form, $\bar{g}_{\text{Na}} = 4.2 m^3 h$, where $m_{\text{ss}}(v) = \alpha_m(v)/(\alpha_m(v) + \beta_m(v))$, $\tau_m = 0.8/(\alpha_m(v) + \beta_m(v))$, $\alpha_m(v) = 0.182(v + 25)/(1 - \exp(-(v + 32.5)/4.5))$, and $\beta_m(v) = 0.124(-v - 32.5)/(1 - \exp((v + 32.5)/4.5))$. $h_{\text{ss}}(v) = 1/(1 + \exp((v + 58)/5))$. $\tau_h = 1/(\alpha_h(v) + \beta_h(v))$, where $\alpha_h(v) = 0.08(v + 40)/(1 - \exp(-(v + 40)/3))$ and $\beta_h(v) = 0.0005(-v - 10)/(1 - \exp((v + 10)/5))$. The parameters for the Na^+ channel model were derived from refs 5 and 44.

The A-type channel models were of the form $\bar{g}_{\text{KA}} = m^4 h$. $m_{\text{ss}}(v) = \alpha_m(v)/(\alpha_m(v) + \beta_m(v))$, where $\alpha_m(v) = -0.01(v + 21.3)/(\exp((v + 21.3)/-35) - 1)$ and $\beta_m(v) = 0.01(v + 21.3)/(\exp((v + 21.3)/35) - 1)$ for proximal channels. $\alpha_m(v) = -0.01(v + 34.4)/(\exp((v + 34.4)/-21) - 1)$ and $\beta_m(v) = 0.01(v + 34.4)/(\exp((v + 34.4)/21) - 1)$ for distal channels. $h_{\text{ss}}(v) = \alpha_h(v)/(\alpha_h(v) + \beta_h(v))$, where $\alpha_h(v) = -0.01(v + 58)/(\exp((v + 58)/-8.2) - 1)$ and $\beta_h(v) = 0.01(v + 58)/(\exp((v + 58)/8.2) - 1)$ for all A-channels. $\tau_m = 0.2 \text{ ms}$. $\tau_h(v) = 5 + (2.6 \text{ ms per mV})(v + 20)/10$ for $v > -20 \text{ mV}$ and 5 ms for $v < -20 \text{ mV}$.

The DR-type channels were of the form $\bar{g}_{\text{KDR}} = m^4$. $m_{\text{ss}}(v) = \alpha_m(v)/(\alpha_m(v) + \beta_m(v))$, where $\alpha_m(v) = -0.0035(v + 30)/(\exp((v + 30)/-13) - 1)$ and $\beta_m(v) = 0.0035(v + 30)/(\exp((v + 30)/13) - 1)$; $\tau_m = 1.8 \text{ ms}$.

Received 11 November 1996; accepted 30 April 1997.

- Johnston, D., Magee, J. C., Colbert, C. M. & Christie, B. R. Active properties of neuronal dendrites. *Annu. Rev. Neurosci.* **19**, 165–186 (1996).
- Yuste, R. & Tank, D. W. Dendritic integration in mammalian neurons, a century after Cajal. *Neuron* **16**, 701–716 (1996).
- Stuart, G. & Sakmann, B. Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. *Nature* **367**, 69–72 (1994).
- Magee, J. C. & Johnston, D. Synaptic activation of voltage-gated channels in the dendrites of hippocampal pyramidal neurons. *Science* **268**, 301–304 (1995).
- Magee, J. C. & Johnston, D. Characterization of single voltage-gated Na^+ and Ca^{2+} channels in apical dendrites of rat CA1 pyramidal neurons. *J. Physiol. (Lond.)* **487**, 67–90 (1995).
- Schwindt, P. C. & Crill, W. E. Amplification of synaptic current by persistent sodium conductance in apical dendrite of neocortical neurons. *J. Neurophysiol.* **74**, 2220–2224 (1995).
- Stuart, G. & Sakmann, B. Amplification of EPSPs by axosomatic sodium channels in neocortical pyramidal neurons. *Neuron* **15**, 1065–1076 (1995).
- Spruston, N., Schiller, Y., Stuart, G. & Sakmann, B. Activity-dependent action potential invasion and calcium influx into hippocampal CA1 dendrites. *Science* **268**, 297–300 (1995).
- Turner, R. W., Meyers, D. E. R., Richardson, T. L. & Barker, J. L. The site for initiation of action potential discharge over the somatodendritic axis of rat hippocampal CA1 pyramidal neurons. *J. Neurosci.* **11**, 2270–2280 (1991).
- Andreasen, M. & Lambert, J. D. C. Regenerative properties of pyramidal cell dendrites in area CA1 of the rat hippocampus. *J. Physiol. (Lond.)* **483**, 421–441 (1995).

- Lipowsky, R., Gillespie, T. & Alzheimer, C. Dendritic Na^+ channels amplify EPSPs in hippocampal CA1 pyramidal cells. *J. Neurophysiol.* **76**, 2181–2191 (1996).
- Hounsgaard, J. & Midtgard, J. Intrinsic determinants of firing pattern in purkinje cells of the turtle cerebellum *in vitro*. *J. Physiol. (Lond.)* **402**, 731–749 (1988).
- Magee, J. C. & Johnston, D. A synaptically-controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* **275**, 209–213 (1997).
- Markram, H. & Tsodyks, M. Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* **382**, 807–810 (1996).
- Christie, B. R., Magee, J. C. & Johnston, D. The role of dendritic action potentials and Ca^{2+} influx in the induction of homosynaptic long-term depression in hippocampal CA1 pyramidal neurons. *Learn. Mem.* **3**, 160–169 (1996).
- Klee, R., Ficker, E. & Heinemann, U. Comparison of voltage-dependent potassium currents in rat pyramidal neurons acutely isolated from hippocampal regions CA1 and CA3. *J. Neurophysiol.* **74**, 1982–1995 (1995).
- Halliwel, J. V., Othman, I. B., Pelchen-Matthews, A. & Dolly, J. O. Central action of dendrotoxin: Selective reduction of a transient potassium conductance in hippocampus and binding to localized acceptors. *Proc. Natl Acad. Sci. USA* **83**, 493–497 (1986).
- Storm, J. F. Temporal integration by a slowly inactivating K^+ current in hippocampal neurons. *Nature* **336**, 379–381 (1988).
- Wu, R.-L. & Barrish, M. E. Two pharmacologically and kinetically distinct transient potassium currents in cultured embryonic mouse hippocampal neurons. *J. Neurosci.* **13**, 2235–2246 (1992).
- Segal, M. & Barker, J. R. Hippocampal neurons in culture: Potassium conductances. *J. Neurophysiol.* **51**, 1409–1433 (1984).
- Segal, M., Rogawski, M. & Barker, J. A transient potassium conductance regulates the excitability of cultured hippocampal and spinal neurons. *J. Neurosci.* **4**, 604–609 (1984).
- Sole, C. K., Zagotta, W. N. & Aldrich, R. W. Single-channel and genetic analyses reveal two distinct A-type potassium channels in *Drosophila*. *Science* **236**, 1094–1097 (1987).
- Sheng, M., Tsaur, M., Jan, Y. N. & Jan, L. Y. Subcellular segregation of two A-type K^+ channel proteins in rat central neurons. *Neuron* **9**, 271–284 (1992).
- Maletic-Savatic, M., Lenn, N. J. & Trimmer, J. S. Differential spatiotemporal expression of K^+ channel polypeptides in rat hippocampal neurons developing *in situ* and *in vitro*. *J. Neurosci.* **15**, 3840–3851 (1995).
- Alonso, G. & Widmer, H. Clustering of Kv 4.2 potassium channels in postsynaptic membrane of rat supraoptic neurons: an ultrastructural study. *Neuroscience* **77**, 617–621 (1997).
- Helmchen, F., Imoto, K. & Sakmann, B. Ca^{2+} buffering and action potential-evoked Ca^{2+} signaling in dendrites of pyramidal neurons. *Biophys. J.* **70**, 1069–1081 (1996).
- Johnston, D. The calcium code. *Biophys. J.* **70**, 1095 (1996).
- Callaway, J. C. & Ross, W. N. Frequency dependent propagation of sodium action potentials in dendrites of hippocampal CA1 pyramidal neurons. *J. Neurophysiol.* **74**, 1395–1403 (1995).
- Jaffe, D. B. *et al.* The spread of Na^+ spikes determines the pattern of dendritic Ca^{2+} entry into hippocampal neurons. *Nature* **374**, 1395–1403 (1995).
- Nicoll, A., Larkman, A. & Blakemore, C. Modulation of EPSP shape and efficacy by intrinsic membrane conductances in rat neocortical pyramidal neurons *in vitro*. *J. Physiol. (Lond.)* **468**, 693–710 (1993).
- Rapp, M., Yarom, Y. & Segev, I. Modeling back propagating action potential in weakly excitable dendrites of neocortical pyramidal cells. *Proc. Natl Acad. Sci. USA* **93**, 11985–11990 (1996).
- Mainen, Z. F., Joerges, J., Huguenard, J. R. & Sejnowski, T. J. A model of spike initiation in neocortical pyramidal neurons. *Neuron* **15**, 1427–1439 (1995).
- Rudy, B. Diversity and ubiquity of K^+ channels. *Neuroscience* **25**, 729–749 (1988).
- Gage, P. W. Activation and modulation of neuronal K^+ channels by GABA. *Trends Neurosci.* **15**, 46–51 (1992).
- Talukder, G. & Harrison, N. L. On the mechanism of modulation of transient outward current in cultured rat hippocampal neurons by di- and trivalent cations. *J. Neurophysiol.* **73**, 73–79 (1995).
- Chen, Q. X. & Wong, R. K. S. Intracellular Ca^{2+} suppressed a transient potassium current in hippocampal neurons. *J. Neurosci.* **11**, 337–343 (1991).
- Serodio, P., Kentros, C. & Rudy, B. Identification of molecular components of A-type channels activating at subthreshold potentials. *J. Neurophysiol.* **72**, 1516–1529 (1994).
- Covarrubias, M., Wei, A., Salkoff, L. & Vyas, T. B. Elimination of rapid potassium channel inactivation by phosphorylation of the inactivation gate. *Neuron* **13**, 1403–1412 (1994).
- Rettig, J. *et al.* Inactivation properties of voltage-gated K^+ channels altered by presence of β -subunit. *Nature* **369**, 289–294 (1994).
- Tsubokawa, H. & Ross, W. N. IPSPs modulate spike backpropagation and associated $[\text{Ca}^{2+}]_i$ changes in the dendrites of hippocampal CA1 pyramidal neurons. *J. Neurophysiol.* **76**, 2896–2906 (1995).
- Yuste, R. & Denk, W. Dendritic spines as basic functional units of neuronal integration. *Nature* **375**, 682–684 (1995).
- Magee, J. C., Avery, R. B., Christie, B. R. & Johnston, D. Dihydropyridine-sensitive, voltage-gated Ca^{2+} channels contribute to the resting intracellular Ca^{2+} concentration of hippocampal CA1 pyramidal neurons. *J. Neurophysiol.* **76**, 3460–3470 (1996).
- Hines, M. in *Neural Systems: Analysis and Modeling* (ed. Eckmann, E.) 127–136 (Kluwer Academic, Norwell, MA, 1993).
- Colbert, C. M. & Johnston, D. Axonal action-potential initiation and Na^+ channel densities in the soma and axon initial segment of subicular pyramidal neurons. *J. Neurosci.* **16**, 6676–6686 (1996).

Acknowledgements. We thank P. Pfaffinger and S. Sinha for comments on the manuscript, and R. Gray for help with computer programming and data analysis. This work was supported by grants from the NIMH, NIH, and the Human Frontiers Science Program.

Correspondence and requests for materials should be addressed to D.J. (e-mail: dan@mossy.bcm.tmc.edu).

The optical counterpart to the γ -ray burst GRB970508

S. G. Djorgovski*, M. R. Metzger*, S. R. Kulkarni*, S. C. Odewahn*, R. R. Gal*, M. A. Pahre*, D. A. Frail†, M. Feroci‡, E. Costa‡ & E. Palazzi§

* Palomar Observatory, California Institute of Technology, Pasadena, California 91125, USA

† National Radio Astronomy Observatory, Socorro, New Mexico 87801, USA

‡ Istituto di Astrofisica Spaziale CNR, 00044 Frascati, Italy

§ Istituto Tecnologie e Studio Radiazioni Extraterrestri CNR, 40129 Bologna, Italy

Understanding the nature of the γ -ray burst phenomenon is one of the outstanding problems of modern astrophysics. The identification of counterparts at optical wavelengths is considered a crucial factor for determining the origin of these events. Here we report the detection and temporal properties of a variable optical source, which has been identified^{1,2} as the counterpart of the X-ray transient associated with the γ -ray burst GRB970508 (ref. 3). The initial optical images were obtained only 5.8 hours after the initial γ -ray burst, after which the optical source was observed to brighten over the next two days before declining in luminosity with a t^{-1} power law. The decline in brightness follows a form predicted by many relativistic fireball models^{4–7} for γ -ray bursts, although the initial rise does not appear to be compatible with the simplest of these models. The observed fluence of the source at visible wavelengths over the period spanned by our observations is $\geq 4.6 \times 10^{-8} \text{ erg cm}^{-2}$, about 3% of the fluence of the γ -ray burst itself.

Since their discovery nearly three decades ago⁸, the understanding of γ -ray bursts (GRBs) has presented a major challenge for astrophysicists, generating an extensive literature^{9–11}. Until recently, progress has been stymied by poor localization at γ -ray wavelengths. It has long been recognized that the key to understanding the nature and origin of GRBs is their detection in some other wavelength regime, preferably in the visible band which has the advantage of spectroscopy, whereby distances could be determined.

The field has been revolutionized with the launch of the BeppoSAX satellite¹², which can provide positions accurate to a few arcmin within hours, and accurate to 1 arcmin (or better) within a fraction of a day. This has already resulted in a major advance: the discovery that GRBs are followed by bright X-ray sources which decay on timescales of the order of a few days (refs 13–15). BeppoSAX can localize these X-ray sources considerably more precisely than their γ -ray positions, to a 45-arcsec radius, or better. This enables rapid follow-up observations at optical and other wavelengths. Another significant new development has been the detection of a fading optical counterpart¹⁶ within the precise localization of the fading X-ray counterpart of GRB970228. No fading optical counterpart was seen¹⁷ in the case of GRB970402, perhaps due to its proximity to the Galactic plane.

Here we report on the optical detection and follow-up observations of the fading X-ray source associated with GRB970508³. The optical counterpart was first reported by Bond¹, and was independently discovered at about the same time in our data². Our observations were conducted on the Palomar 200-inch Hale telescope (hereafter P200) and the Palomar 60-inch telescope (hereafter P60). On the P200 we used a prime-focus charge-coupled device (CCD) imager, which has a field of view of 13.6 arcmin with 0.40-arcsec pixels. On the P60 we used a CCD camera attached to the Cassegrain focus, with a field of 12.8 arcmin and 0.372-arcsec pixels. All observations were in the Gunn gri photometric system.

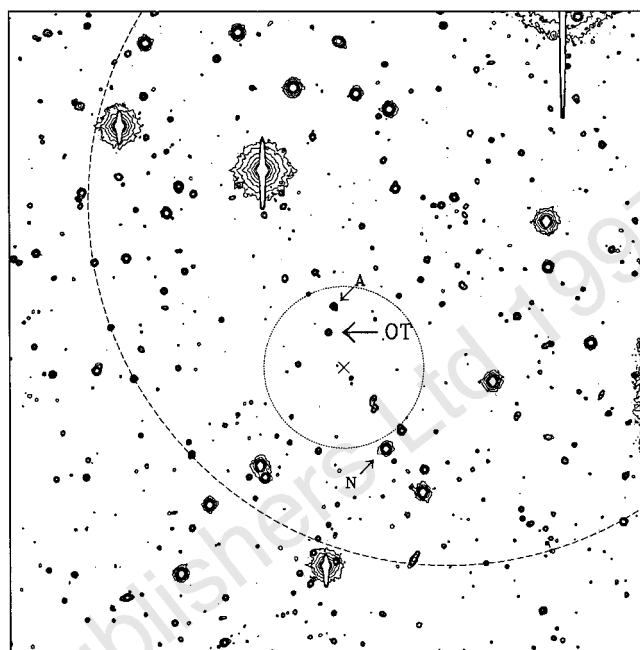


Figure 1 Finding chart for the field of GRB970508, from an r-band image obtained at Palomar on 1997 May 10 UT. The field shown is 320-arcsec square, with north at the top and east to the left. The proposed optical counterpart of the burst (OT) is indicated (see text). We also indicate two nearby offset stars, labelled 'A' (2.7 arcsec west and 13.0 arcsec north of the OT, $r = 19.85$ mag) and 'N' (28.9 arcsec west and 58.3 arcsec south of the OT, $r = 17.4$ mag). The dotted circle centred on the cross, with a 40-arcsec radius, indicates the error circle of the X-ray transient¹⁵. The dashed circle, with a radius of 3 arcmin, indicates the final error circle of the GRB¹⁸.

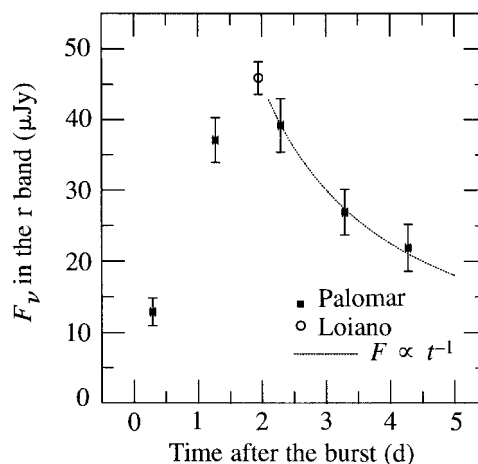


Figure 2 Optical light curve in the r band, as a function of time starting at May 8.904 UT, from the Palomar data listed in Table 1 (filled squares). The data have been corrected for the Galactic extinction and converted to specific fluxes as described in the text. These numbers supersede our preliminary measurements²⁸. We also include a Loiano Observatory measurement²⁹ of $R_C = 19.78 \pm 0.05$ mag (in the Kron-Cousins system) at the mean epoch of May 10.85, close to the peak of the light curve (empty circle). After correcting for the extinction and using the R_C zero-point calibration³⁰, we derive $F_r = 45.9 \mu\text{Jy}$ at the effective wavelength $\lambda_{\text{eff}} \approx 641$ nm, close to the centre of the r band. We do not use other magnitudes reported in the IAU Circulars so far, owing to the lack of information about the photometric systems used, photometric conditions, and so on. The decline in brightness seen after 10 May is consistent with a power-law $F_r \propto t^{-1}$, with time counting from the GRB event, shown as the dotted line drawn through our 11 May data point.

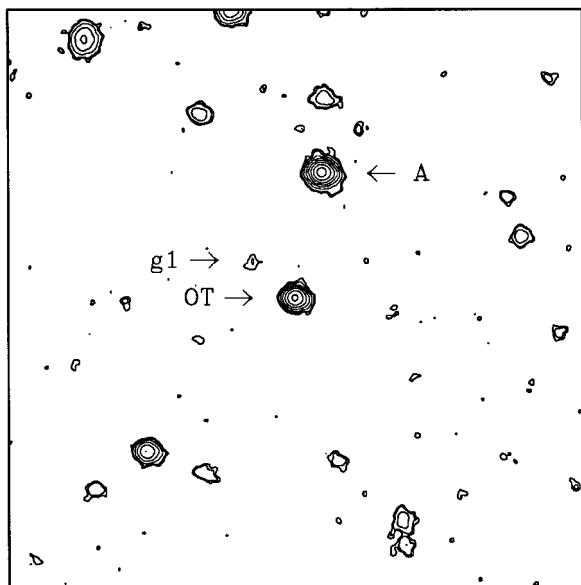


Figure 3 A section of a deep r-band image (average of our P200 observations on 9 and 10 May), centred on the OT. The field shown is 1-arcmin square, with north at the top, and east to the left. A faint, ~ 24.5 mag galaxy ('g1') is seen 4.3 arcsec east and 3.5 arcsec north of the OT. It may be associated with the host galaxy of the OT itself, or with one of the absorption line systems seen in its spectrum^{26,27}. The feature labelled 'A' is the offset star marked in Fig. 1.

The first set of observations were conducted on P200 starting only 5.8 hours after the burst. Multiple exposures in the Gunn r band were obtained in several pointings to cover the large 10-arcmin BeppoSAX Wide Field Camera error circle³. Typical limiting magnitudes were $r \approx 23.5$ mag for the individual 5-minute exposures, and $r \approx 24.5$ mag for the average stack of them. No source brighter than $r \approx 22$ mag was seen to vary by more than 20% over the baselines of about 1–2 hours in these data. On the next night, 1997 May 10 UT, images were obtained but this time centred on the revised GRB position¹⁸, with an error radius of 3 arcmin. Some images in the Gunn g and i bands were also obtained on this night.

On the same night, 10 May, gri observations were also made at the P60 and calibrated using standard stars¹⁹. These data were used to establish a grid of secondary photometric stars in the field. The calibration observations were obtained under photometric conditions. The observations of this night define the photometric zero-points, which we estimate to be good to $\Delta g \leq 0.08$ mag, $\Delta r \leq 0.06$ mag, and $\Delta i \leq 0.08$ mag (conservative errors). All the magnitudes are referred to this grid of secondary stars.

Subsequent observations were obtained at the P60 on 1997 May 11.149 to 11.238 UT, and May 12.156 to 12.235 UT, and at the P200 on 1997 May 13.164 to 13.194 UT. Some of these observations were obtained in non-photometric conditions, but because we use relative photometry from a grid of secondary photometric stars in the field, corrected magnitudes can still be obtained. From the scatter of magnitude differences for different nights, we estimate that these offsets are determined to 0.02–0.03 mag accuracy.

The comparison of P200 data between 9 and 10 May indicates several possible variable sources over the entire 3-arcmin-radius GRB error circle. Variable candidates were identified in an automated fashion using a combination of the photometry program DoPHOT²⁰ and an object cataloguing package²¹. However, only one

Table 1 Photometric observations of GRB970508

Date (UT)	Telescope	Band	Magnitude	F_ν (μ Jy)
May 9.195	P200	r	21.35 ± 0.15	12.9
May 10.199	P200	g	20.25 ± 0.10	29.5
May 10.178	P200	r	20.20 ± 0.09	37.1
May 10.238	P200	i	20.22 ± 0.10	38.8
May 11.198	P60	r	20.14 ± 0.10	39.2
May 12.195	P60	r	20.55 ± 0.12	26.9
May 13.179	P200	r	20.77 ± 0.15	21.9

P200, Palomar 200-inch Hale telescope; P60, Palomar 60-inch telescope.

of them—the highest-amplitude object—is within the X-ray transient source error circle¹⁵. We measure for its position in the US Naval Observatory's A1.0 reference system: right ascension (RA) 06 h 53 min 49.43 s, declination (dec.) $+79^\circ 16' 19.6''$ (J2000) with uncertainties of ~ 1 arcsec. The same optical transient (hereafter OT) was reported by Bond¹. We measure brightening of 1.15 ± 0.10 mag, from $r = 21.35 \pm 0.15$ mag on May 9.195 UT, to $r = 20.20 \pm 0.09$ mag on May 10.178 UT (mean epochs). Given that the source is within the 45-arcsec (radius) error circle of the X-ray transient, and given its peculiar light curve, we conclude that this is the probably optical counterpart of the X-ray transient, which in turn is most likely to be associated with GRB970508. A finding chart for the object is shown in Fig. 1.

From the Infra-Red Astronomy Satellite (IRAS) 100- μ m cirrus flux in this direction²² we derive for the Galactic extinction $A_B = 0.11$ mag. Using the interpolation of the extinction curve²³, we derive extinction in the gri bands of $A_g = 0.09$ mag, $A_r = 0.07$ mag and $A_i = 0.05$ mag. These corrections are applied to all of our measurements listed below.

Table 1 summarizes the photometric results on the OT, after applying the extinction corrections above. Using the flux-standard stars²⁴, we can convert these magnitudes to specific fluxes; we assume that $r = 0$ mag corresponds to the specific flux $F_\nu = 4,460$ Jy. The effective wavelength of the Gunn r band is $\lambda_{\text{eff},r} \approx 665$ nm. The light curve is shown in Fig. 2.

On May 10.178, the extinction-corrected colours of the source were $(g - r) = +0.07$ mag and $(r - i) = -0.09$ mag. To bring all measurements to a common epoch, we assumed a brightening rate of 0.05 mag h^{-1} in all three bands. Assuming that $g = 0$ mag corresponds to $F_\nu = 3,720$ Jy at $\lambda_{\text{eff},g} \approx 520$ nm, and $i = 0$ mag to $F_\nu = 4,750$ Jy at $\lambda_{\text{eff},i} \approx 790$ nm, we can fit the gri fluxes with a power-law continuum of the form $F_\nu \propto \nu^\alpha$. We get $\alpha = -0.65 \pm 0.3$, in excellent agreement with the slope estimated from the nearly simultaneous spectroscopy²⁵.

The light curve shows an initial rise and then a decline (starting on 11 May) consistent with a t^{-1} (or, more precisely, $t^{-0.9 \pm 0.1}$) dependence, where t is the time since the burst. This is indeed predicted by the fireball model^{4–7}. This is a generic model in which a large amount of energy is rapidly released in a small volume. Initially, the deposited energy is optically thick to pair creation ($\gamma\gamma \rightarrow e^+e^-$), and consequently radiation pressure drives a rapid expansion in a fireball stage. The expansion cools the electrons and all the energy appears as kinetic energy of the baryons in a bulk outflow phase. When this shell meets the interstellar medium, the resulting shock produces broad-band emission. In this model, the initial γ -ray burst is expected to be followed by X-ray, optical and radio outbursts^{7,26,31}, as indeed observed³² for GRB970508.

However, we also see a relatively slow rise from 9 to 10 May. If the rise was exponential between these two points, the e -folding time (that is, the time in which the flux would increase by a factor of $e = 2.718$...) would be ~ 22 hours in the observed reference frame. This is not predicted in the simplest fireball model, where the onset of the optical emission should be very rapid⁷. But, inverse Compton

cooling of relativistic electrons in the early stages of expansion could account for such a behaviour³³.

We integrate the r-band light curve over the time interval spanned by our observations to obtain a total fluence. We interpolate linearly in time between the measured magnitudes, and integrate the product νF_ν , where ν is the effective frequency in the r band, between the first and the last of our observations to obtain the total fluence in the visible (observed red) light $I_r > 4.6 \times 10^{-8} \text{ erg cm}^{-2}$. This is a lower limit, as we do not include the unknown portions of the light curve outside our measurements, and we do not sample well the portion near maximum light. For comparison, the γ -ray burst had a fluence of $I_\gamma \approx (1.8 \pm 0.3) \times 10^{-6} \text{ erg cm}^{-2}$, derived from BeppoSAX observations in the energy band 40–700 keV. We thus conclude that at least 3% of the total GRB energy is emitted as visible light.

Figure 3 shows a stack of all of our P200 images from 9 and 10 May, centred on the source. A faint, blue galaxy (labelled 'g1') with $g \approx 24.5 \pm 0.4 \text{ mag}$ and $r \approx 24.8 \pm 0.4 \text{ mag}$ is clearly detected 4.3 arcsec east and 3.5 arcsec north of the OT. The two may be physically associated, or the galaxy could be responsible for some of the absorption lines seen in the source spectrum reported by Metzger *et al.*^{25,27}. Its magnitude is typical of the general field galaxy population at $z \approx 0.8$, and its blue colour is suggestive of active star formation. $\square$

Received 21 May; accepted 28 May 1997.

1. Bond, H. E. *IAU Circ. No. 6654* (1997).
2. Djorgovski, S. *et al. IAU Circ. No. 6655* (1997).
3. Costa, E. *et al. IAU Circ. No. 6649* (1997).
4. Paczyński, B. Gamma-ray bursters at cosmological distances. *Astrophys. J.* **308**, L43–L45 (1986).
5. Goodman, J. Are gamma-ray bursts optically thick? *Astrophys. J.* **308**, L47–L50 (1986).
6. Mészáros, P. & Rees, M. J. Optical and long-wavelength afterglow from gamma-ray bursts. *Astrophys. J.* **476**, 232–237 (1997).
7. Wijers, R. A. M. J., Rees, M. J. & Mészáros, P. Shocked by GRB 970228: the afterglow of a cosmological fireball. *Mon. Not. R. Astron. Soc.* (in the press); preprint <http://xxx.lanl.gov/astro-ph/9704153>.
8. Kleban, R. W., Strong, I. B. & Olson, R. A. Observations of gamma-ray bursts of cosmic origin. *Astrophys. J.* **182**, L85–L88 (1973).
9. Fishman, G. J. & Meegan, C. A. Gamma-ray bursts. *Annu. Rev. Astron. Astrophys.* **33**, 415–458 (1995).
10. Lamb, D. Q. The distance scale to gamma-ray bursts. *Publ. Astron. Soc. Pacif.* **107**, 1152–1166 (1995).
11. Paczyński, B. How far away are gamma-ray bursters? *Publ. Astron. Soc. Pacif.* **107**, 1167–1175 (1995).
12. Boella, G. *et al.* BeppoSAX, the wide band mission for X-ray astronomy. *Astron. Astrophys. Suppl. Ser.* **122**, 299–399 (1997).
13. Costa, E. *et al. IAU Circ. No. 6572*; (1997); Discovery of an X-ray afterglow associated with the γ -ray burst of 28 February 1997. *Nature* **387**, 783–785 (1997).
14. Piro, L. *et al. IAU Circ. No. 6617* (1997).
15. Piro, L. *et al. IAU Circ. No. 6656* (1997).
16. van Paradijs, J. *et al.* Transient optical emission from the error box of the gamma-ray burst of 28 February 1997. *Nature* **386**, 686–689 (1997).
17. Harrison, T. *et al. IAU Circ. No. 6632* (1997).
18. Heise, J. *et al. IAU Circ. No. 6654* (1997).
19. Kent, S. M. Photometry of stars in the uvgr system. *Publ. Astron. Soc. Pacif.* **97**, 165–174 (1985).
20. Schechter, P. L., Mateo, M. & Saha, A. DoPHOT, a CCD photometry program—description and tests. *Publ. Astron. Soc. Pacif.* **105**, 1342–1353 (1993).
21. Metzger, M. R. Constraints on the shape and size of the Galactic gravitational potential. Thesis, Massachusetts Inst. Technol. (1994).
22. Laureijs, R. J., Helou, G. & Clark, F. O. *The First Symp. on the Infrared Cirrus and Diffuse Interstellar Clouds* (eds Cutri, R. & Latter, W.) 133–136 (Conf. Ser. 58, Astron. Soc. Pacif., Provo, UT, 1994).
23. Seaton, M. J. Interstellar extinction in the UV. *Mon. Not. R. Astron. Soc.* **187**, 73P–76P (1979).
24. Oke, J. B. & Gunn, J. E. Secondary stars for absolute spectrophotometry. *Astrophys. J.* **266**, 713–717 (1983).
25. Metzger, M. R. *et al.* Spectral constraints on the redshift of the optical counterpart to the γ -ray burst of 8 May 1997. *Nature* **387**, 878–880 (1997).
26. Mészáros, P. & Rees, M. J. Gamma-ray burst models. *Astrophys. J.* **397**, 570–575 (1992).
27. Metzger, M. R. *et al. IAU Circ. No. 6655* (1997).
28. Djorgovski, S. *et al. IAU Circ. No. 6658* (1997).
29. Mignoli, M. *et al. IAU Circ. No. 6661* (1997).
30. Bessel, M. UBVRi passbands. *Publ. Astron. Soc. Pacif.* **102**, 1181–1199 (1990).
31. Paczyński, B. & Rhoads, J. E. Radio transients from gamma-ray bursters. *Astrophys. J.* **418**, L5–L8 (1993).
32. Frail, D. A. & Kulkarni, S. R. *IAU Circ. No. 6662* (1997).
33. Waxman, E. GRB after-glow: supporting the cosmological fireball model, constraining parameters, and making predictions. *Astrophys. J.* (in the press); preprint <http://xxx.lanl.gov/astro-ph/9704116> (1997).

Acknowledgements. We thank the staff of Palomar Observatory for their assistance during our observing runs, and the entire BeppoSAX team for their efforts in providing accurate burst positions as rapidly as possible. This work was supported in part by the NSF, NASA, the Bressler Foundation, the Norris Foundation and Caltech.

Correspondence should be addressed to S.G.D. (e-mail: george@oracle.caltech.edu).

Spectral constraints on the redshift of the optical counterpart to the γ -ray burst of 8 May 1997

M. R. Metzger*, S. G. Djorgovski*, S. R. Kulkarni*, C. C. Steidel*, K. L. Adelberger*, D. A. Frail†, E. Costa‡ & F. Frontera§

* Palomar Observatory, 105–24, California Institute of Technology, Pasadena, California 91125, USA

† National Radio Astronomy Observatory, Socorro, New Mexico 87801, USA

‡ Istituto di Astrofisica Spaziale CNR, 00044 Frascati, Italy

§ Dipartimento di Fisica, Università di Ferrara and Istituto TESRE-CNR, 40129 Bologna, Italy

Brief, intense bursts of γ -rays occur approximately daily from random directions in space, but their origin has remained unknown since their initial detection almost 25 years ago¹. Arguments based on their observed isotropy and apparent brightness distribution² are not sufficient to constrain the location of the bursts to a local³ or cosmological origin⁴. The recent detection of a counterpart to a γ -ray burst at other wavelengths^{5,6} has therefore raised the hope that the sources of these energetic events might soon be revealed. Here we report spectroscopic observations of the possible optical counterpart^{7,8} to the γ -ray burst GRB970508. The spectrum is mostly featureless, except for a few prominent absorption lines which we attribute to the presence of an absorption system along the line of sight at redshift $z = 0.835$. Coupled with the absence of Lyman- α forest features in the spectra, our results imply that the optical transient lies at $0.835 \leq z \leq 2.3$. If the optical transient is indeed the counterpart of GRB970508, our results provide the first direct limits on the distance to a γ -ray burst, confirming that at least some of these events lie at cosmological distances, and are thus highly energetic.

On 8 May 1997 UT, a moderate-fluence classical γ -ray burst (GRB970508) was detected by instruments aboard the Italian–Dutch satellite BeppoSAX⁹; the burst was localized initially to an error region of 5-arcmin radius¹⁰ and later to a region of 3-arcmin radius¹¹. A potential optical counterpart was identified within two days^{7,8}, which we refer to as OT J065349+79163 (here OT stands for optical transient). Interest in this object was heightened when the appearance of a bright X-ray source was reported¹² that was not seen in a previous X-ray all-sky survey, and included OT J065349+79163 in the 45-arcsec-radius error circle. The presence of a new and bright X-ray source in a γ -ray burst (GRB) error circle has been seen in the recent three GRBs observed by BeppoSAX. In the first such case, a decaying optical source⁵ was also seen and these authors suggested that such fading (X-ray and optical) sources are the afterglow of γ -ray bursts. It has been suggested¹³ that because OT J065349+79163 exhibits unusual variability at optical and X-ray wavelengths, it represents a similar optical afterglow to GRB970508.

We obtained spectra of OT J065349+79163 with the Keck II 10-m telescope on 11 May 1997 UT, using the Low Resolution Imaging Spectrograph¹⁴ (LRIS) with a $2,048 \times 2,048$ pixel CCD (charge-coupled device). Starting at 05:44 UTC, a sequence of three 10-minute spectra were obtained with a 1.0-arcsecond-wide slit oriented along the direction of the atmospheric dispersion. The spectrograph was configured with a 300 lines mm^{-1} grating blazed at 5,000 Å, covering the region 3,850–8,550 Å. We obtained two further spectra of 10 minutes each on 12 May 1997 UT. Calibration Hg–Kr–Ar and flat lamp spectra were taken at the end of each exposure sequence. Owing to the extremely low elevation angle of

Table 1 OT J065349+79163 absorption lines

λ_{vac} (Å)	Unc.	W_{λ} (Å)	Unc.	λ_{rest} Å	z	Assignment
4,302.5	1.8	1.3	0.3	2,344.2	0.8354(8)	Fe II
4,359.7	1.4	1.3	0.3	2,374.5	0.8360(6)	Fe II
4,372.2	1.5	1.4	0.3	2,382.8	0.8349(6)	Fe II
4,746.7	1.7	1.0	0.4	2,586.7	0.8350(7)	Fe II
4,769.7	1.3	2.3	0.2	2,600.2	0.8344(5)	Fe II
4,941.1	1.5	1.3	0.3	2,796.4	0.7670(5)	Mg II
4,953.9	1.5	1.0	0.4	2,803.5	0.7670(5)	Mg II
5,130.4	1.1	2.7	0.2	2,796.4	0.8346(4)	Mg II
5,144.0	1.1	3.0	0.2	2,803.5	0.8348(4)	Mg II
5,232.6	1.3	1.8	0.2	2,853.0	0.8341(5)	Mg I

Table gives measured parameters for identified absorption lines in OT J065349+79163 and the inferred redshift of each feature. λ_{vac} is the measured wavelength of each line, corrected to vacuum, and the following column is the uncertainty (in Å); W_{λ} is the observed (not rest frame) equivalent width of the line in Å, along with the corresponding uncertainty; the last three columns list the assigned physical absorption for each line, with rest vacuum wavelength (λ_{rest}), implied redshift, and element/ionization state.

the telescope during the observations ($\sim 25^\circ$) and a problem with the Keck II lower shutter, approximately half of the telescope aperture was blocked during the exposures. Instrumental sensitivity was calibrated by an observation of the standard star¹⁵ BD+284211, but owing to an approximate correction for the occulted aperture the absolute calibration is only rough.

The resulting combined spectra from 11 May is shown in Fig. 1a. It is customary to represent the spectral flux density (F_ν) of a non-thermal source by a power law, $F_\nu \propto \nu^\alpha$; here ν is the frequency. The optical index computed from the spectrum is $\alpha_o = -0.9 \pm 0.3$. The large uncertainty is due to the uncertain correction for atmospheric extinction in the blue region of the spectrum at the large zenith angles of our observations.

Several absorption features are evident; the strongest, near 7,600 and 6,870 Å, are due to telluric O₂. In the region between 4,300 and 5,300 Å (Fig. 1b), there are several significant absorption features¹⁶ that we identify. The identifications were made based on Mg II doublet (5,129 and 5,143 Å) line ratios, and assigning further rough identification of other metal lines based on wavelength ratios between these and the Mg doublet. Table 1 shows the lines identified in the spectrum; independent redshifts are computed from each line. This reveals a relatively strong¹⁷ metal line absorption system at $z = 0.8349 \pm 0.0002$, and a weaker Mg II system at $z = 0.768$. The eight lines present in the strong absorption system make the redshift assignment unambiguous. The continuum source is either more distant and absorbed by a gas cloud at this redshift, or perhaps is located physically within the cloud, but the absorption places a firm lower limit to the redshift of the source, $z \geq 0.835$.

Such absorption systems are commonly seen in the spectra of high-redshift quasi-stellar objects (QSOs)¹⁷. An imaging study of such systems¹⁸ reveals that most are associated with normal galaxies close to the line of sight to the QSO. An analysis of these systems¹⁹ at redshift similar to the system we identify in OT J065349+79163 indicates a correlation (with significant scatter) between line equivalent width and impact parameter. As the absorption we see should be similar to QSO systems, we expect that deep images (perhaps taken after the transient fades) would reveal a galaxy responsible for this absorption system, though it is difficult to predict its brightness or separation from the transient. A hint of such an object has already been suggested²⁰. Note that as the OT was far brighter than any other nearby object, any contamination of the spectrum is negligible and thus the OT features were a physical absorption.

At these redshifts, the number of Mg II absorption systems with rest equivalent widths $W_\lambda > 0.3$ Å per unit redshift is of the order of unity¹⁷. Detection of one or two such absorption systems in our spectrum is thus not unusual. However, the ratio of line strengths (Mg I/Mg II) seems unusually high, and combined with the high

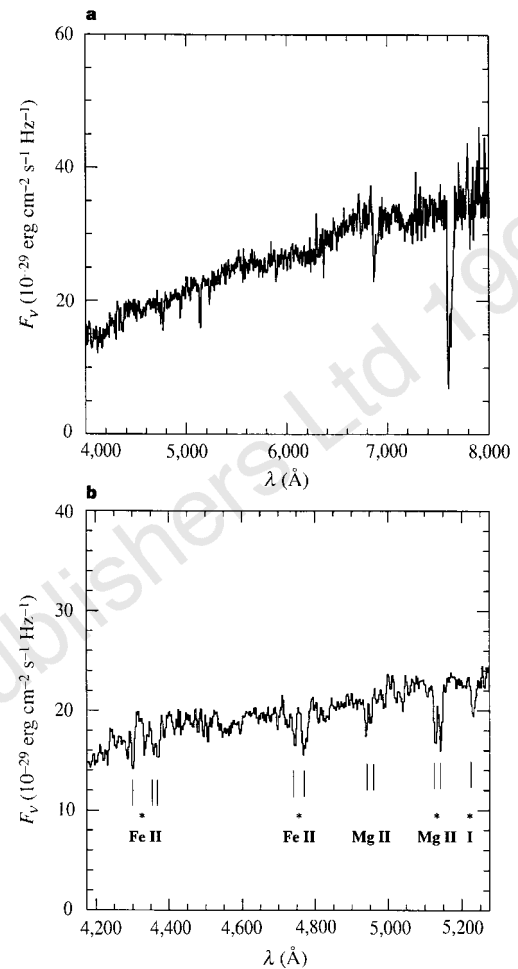


Figure 1 The spectrum of the optical variable. **a**, Full spectrum; **b**, expansion of a limited region, with strong absorption lines and identifications indicated. The lines marked with an asterisk are identified with an absorption system at redshift $z = 0.835$, the others at $z = 0.767$. The spectrum has been smoothed with a three-pixel boxcar filter. A few additional weak features (not shown) have also been tentatively identified with the $z = 0.767$ system. F_ν is the flux density, and λ is the wavelength in Å.

strength of the Mg II absorption system provides some evidence for a dense foreground interstellar medium. This implies either a small impact parameter¹⁹, or, more likely, that the $z = 0.835$ system is due to the GRB host galaxy itself. We can also place an approximate upper limit to the source redshift from the absence of apparent Lyman- α absorption features in our spectra. The short-wavelength limit of our data corresponds to $z_{\text{Ly}\alpha} \approx 2.3$. In addition to the lack of individual lines, the mean observed continuum decrement at this redshift is^{21,22} $D_A \approx 0.1$ – 0.2 , and it increases with redshift. If present, such a continuum drop should be detected in our data for wavelength $\lambda > 4,000$ Å. We can thus place an approximate upper limit to the source redshift of $z \lesssim 2.3$.

One might ask whether from current observations we should expect to see a host galaxy for the burst, if such a galaxy were present. If we assume a minimum redshift of $z = 0.835$ in a standard Friedmann cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_0 = 0.2$, the luminosity distance is $1.49 \times 10^{28} \text{ cm}$. The B band would be redshifted just slightly past the Gunn i band, and for observations¹³ made on 10 May UT, the observed flux in the redshifted B band is $\sim 39 \mu\text{Jy}$. For the assumed redshift and cosmology, this implies an

absolute magnitude of $M_B \approx -22.6$ mag, or a lower limit for $L_B \approx 7L_*$, where L_* is a characteristic galaxy luminosity³³. Thus, OT J065349+79163 is still significantly outshining any host galaxy for the most probable host luminosity range. The properties of the Mg absorption lead us to expect that once the OT has faded, the galaxy could be identified optically.

Taken together, the source's compact optical appearance, a featureless continuum, X-ray emission and high redshift suggest a possible classification (independent of a burst event) as a BL Lac object²³. One of the known characteristics of BL Lac objects is their variability from radio to γ -ray wavelengths. We now evaluate the a posteriori probability that we might be seeing a BL Lac object by random coincidence with the γ -ray error box. The surface density of BL Lac objects with Rosat X-ray flux $f_X \leq 10^{-12}$ erg cm⁻² s⁻¹ is not very well known, but there are indications that this distribution is quite flat at low flux densities. A simple extrapolation for the expected number of BL Lacs with $f_X > 6 \times 10^{-13}$ erg cm⁻² s⁻¹ is²⁴ 0.03 per square degree. Thus the probability of finding a BL Lac object within the 3-arcmin-radius localization region is $\sim 2 \times 10^{-4}$. The amplitude of the variability detected in the counterpart¹³ over a few days is also larger than has been observed in studies of BL Lac object variability^{25,26}. Although we cannot completely exclude the possibility that OT J065349+79163 is a chance coincidence of a BL Lac object with the GRB error circle, the probability of finding a random BL Lac object which also exhibits variability that is temporally correlated with a γ -ray burst is quite small. Thus we conclude that the OT is probably associated with GRB970508, regardless of classification, though the strongest constraints naturally come from higher-energy emission.

The high redshift of OT J065349+79163, its featureless spectrum and slowly decaying optical flux are consistent with the so-called fireball models for cosmological bursts^{27–29}, which are efficient at emitting γ -rays and produce power-law spectral energy distributions. The fluence³⁰ of GRB970508 in the energy range 20–1,000 keV was 3×10^{-6} erg cm⁻², and at the minimum redshift implied for OT J065349+79163, this burst would have a total γ -ray energy of 7×10^{51} erg (assuming isotropic emission). This falls in the general range of typical γ -ray burst energies from various cosmological models^{31,32}.

The remarkable progress in detecting X-ray and optical counterparts to GRBs has been made possible only by rapid localization of the burst by BeppoSAX and prompt dissemination of the coordinates by the BeppoSAX team. Further progress in understanding GRBs requires many more optical counterparts to be identified. It is clear from experience of the first two optical counterparts that, in order to obtain the critical data, the counterparts must be discovered and followed up spectroscopically within a few days. It now seems that an understanding of the physical mechanisms behind γ -ray bursts is within reach.

Received 21 May; accepted 3 June 1997.

- Klebesadel, R. W., Strong, I. B. & Olsen, R. A. Observations of gamma-ray bursts of cosmic origin. *Astrophys. J.* **182**, L85–L88 (1973).
- Meegan, C. A. *et al.* Spatial distribution of gamma-ray bursts observed by BATSE. *Nature* **355**, 143–145 (1992).
- Lamb, D. Q. The distance scale to gamma-ray bursts. *Publ. Astron. Soc. Pacif.* **107**, 1152–1166 (1995).
- Paczynski, B. How far away are gamma-ray bursters? *Publ. Astron. Soc. Pacif.* **107**, 1167–1175 (1995).
- van Paradijs, J. *et al.* Transient optical emission from the error box of the γ -ray burst of 28 February 1997. *Nature* **386**, 686–689 (1997).
- Costa E. *et al.* *IAU Circ.* No. 6576 (1997).
- Bond, H. E. *IAU Circ.* No. 6654 (1997).
- Djorgovski, S. *et al.* *IAU Circ.* No. 6655 (1997).
- Boella, G. *et al.* BeppoSAX, the wide band mission for X-ray astronomy. *Astron. Astrophys. Suppl.* **122**, 299 (1997).
- Costa, E. *et al.* *IAU Circ.* No. 6649 (1997).
- Heise, J. *et al.* *IAU Circ.* No. 6654 (1997).
- Piro, L. *et al.* *IAU Circ.* No. 6656 (1997).
- Djorgovski, S. G. *et al.* The optical counterpart to the γ -ray burst GRB970508. *Nature* **387**, 876–878 (1997).
- Oke, J. B. *et al.* The Keck low-resolution imaging spectrometer. *Publ. Astron. Soc. Pacif.* **107**, 375–385 (1995).
- Massey, P., Strobel, K., Barnes, J. V. & Anderson, E. Spectrophotometric standards. *Astrophys. J.* **328**, 315–333 (1988).

- Metzger, M. R. *et al.* *IAU Circ.* No. 6655 (1997).
- Steidel, C. C. & Sargent, W. L. W. Mg II absorption in the spectra of 103 QSOs: implications for the evolution of gas in high-redshift galaxies. *Astrophys. J. Suppl. Ser.* **80**, 1–108 (1992).
- Steidel, C. C., Dickinson, M. & Persson, S. E. Field galaxy evolution since z approximately 1 from a sample of QSO absorption-selected galaxies. *Astrophys. J.* **437**, L75–L78 (1994).
- Steidel, C. C. in *QSO Absorption Lines* (ed. Meylan, G.) 139–152 (Springer, Berlin, 1995).
- Djorgovski, S. *et al.* *IAU Circ.* No. 6660 (1997).
- Oke, J. B. & Korycansky, D. G. Absolute spectrophotometry of very large redshift quasars. *Astrophys. J.* **255**, 11–19 (1982).
- Zuo, L. & Lu, L. Measurements of D_A for a large QSO sample and determination of evolution of Lyman-alpha clouds. *Astrophys. J.* **418**, 601–616 (1993).
- Stoeck, J. T. *et al.* The Einstein Observatory extended medium-sensitivity survey. II – The optical identifications. *Astrophys. J. Suppl. Ser.* **76**, 813–874 (1991).
- Nass, P. *et al.* BL Lacertae objects in the ROSAT All-Sky Survey: new objects and comparison of different search techniques. *Astron. Astrophys.* **309**, 419–430 (1996).
- Miller, H. R. & Noble, J. C. The microvariability of blazars and related AGN. 17–29 (ASP Conf. Ser. 110, Astron. Soc. Pacif., 1996).
- Heidt, J. & Wagner, S. J. Statistics of optical intraday variability in a complete sample of radio-selected BL Lacertae objects. *Astron. Astrophys.* **305**, 42–52 (1995).
- Nézsáros, P. & Rees, M. J. Gamma-ray bursts; multiwaveband spectral predictions for blast wave models. *Astrophys. J.* **418**, L59–L62 (1993).
- Paczynski, B. & Rhoads, J. E. Radio transients from gamma-ray bursters. *Astrophys. J.* **418**, L5–L8 (1993).
- Waxman, E. Gamma-ray burst after-glow: confirming the cosmological fireball model. *Astrophys. J.* (submitted).
- Kouveliotou, C. *et al.* *IAU Circ.* No. 6660 (1997).
- Fenimore, E. E. *et al.* The intrinsic luminosity of gamma-ray bursts and their host galaxies. *Nature* **366**, 40–42 (1993).
- Fenimore, E. E. & Bloom, J. S. Determination of distance from time dilation of cosmological gamma-ray bursts. *Astrophys. J.* **453**, 25–36 (1995).
- Schachter, P. An analytic expression for the luminosity function of galaxies. *Astrophys. J.* **203**, 297–306 (1976).

Acknowledgements. We thank the BeppoSAX team for their efforts in disseminating information to observers for rapid identification of an optical transient, making this work possible. We thank W. Sargent and M. Pahre for discussions, and the Keck Observatory staff for assistance at the telescope. This work is based on observations obtained at the W. M. Keck Observatory, which is jointly operated by the California Institute of Technology and the University of California. M.R.M. was supported by Caltech; S.R.K. was supported by NASA and NSF; S.G.D. was supported by NSF and the Bressler Foundation; C.C.S. was supported by NSF and the Sloan Foundation.

Correspondence should be addressed to M.R.M. (e-mail: mrm@astro.caltech.edu).

Self-trapping of incoherent white light

Matthew Mitchell & Mordechai Segev

Electrical Engineering Department, Center for Photonics and Optoelectronic Materials and Princeton Materials Institute, Princeton University, Princeton, New Jersey 08544, USA

Optical pulses—wave-packets—propagating in a linear medium have a natural tendency to broaden in time (dispersion) and space (diffraction). Such broadening can be eliminated in a nonlinear medium that modifies its refractive index in the presence of light in such a way that dispersion or diffraction effects are counteracted by light-induced lensing^{1,2}. This can allow short pulses to propagate without changing their shape^{2,3}, and the ‘self-trapping’ of narrow optical beams¹ whereby a beam of light induces a waveguide in the host medium and guides itself in this waveguide, thus propagating without diffraction⁴. Self-trapped pulses in space and time have been investigated extensively in many physical systems and, as a consequence of their particle-like behaviour, are known as ‘solitons’ (ref. 5). Previous studies of this phenomenon in various nonlinear media^{6–12} have involved coherent light, the one exception being our demonstration¹³ of self-trapping of an optical beam that exhibited partial spatial incoherence. Here we report the observation of self-trapping of a white-light beam from an incandescent source. Self-trapping occurs in both dimensions transverse to the beam when diffraction effects are balanced exactly by self-focusing in the host photorefractive medium. To the best of our knowledge, this is the first observation of self-trapping for any wave-packet that is both temporally and spatially incoherent.

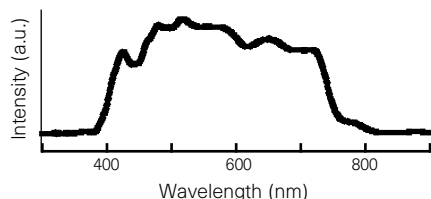


Figure 1 Spectrum of the incoherent white-light beam incident on the photorefractive crystal (a.u., arbitrary units).

The goal is to achieve self-trapping of an incoherent beam (wavepacket); that is, we wish to trap the time-averaged envelope made by a rapidly changing multi-mode broad-band optical field. At any given time, the beam contains many 'speckles': a random distribution of bright and dark patches caused by the randomly varying phase in space. To achieve self-trapping of the envelope, it is necessary to counteract the diffraction of the quickly changing speckled optical beam. The key issue is to find a nonlinear medium that responds on a timescale much longer than the rate of change of the speckle pattern of the beam. The reason for this is as follows: if an instantaneous optical nonlinearity was used, then the medium would respond to the instantaneous 'speckled' beam. Each speckle would form a small 'positive lens' and would capture a small fraction of the beam. These bright-dark features on the beam change very fast throughout propagation and these tiny induced-waveguides intersect and cross each other in a random manner. The net effect would be beam breakup into small fragments and self-trapping of the beam's envelope would not occur. For an incoherent beam to self-trap, it is necessary that the beam induces a smooth waveguide that will guide its rapidly changing intensity at every instant. A non-instantaneous nonlinearity allows this to occur, as it reacts to a time-averaged intensity which is temporally and spatially smooth if averaged over a long enough time period. When this self-induced waveguide guides the rapidly changing beam, self-trapping is achieved.

Photorefractive crystals are a convenient choice of such a nonlinear medium, because their response time is fully controlled by the intensity of the beam, and can be made (with low intensities) much longer than the rate of the rapid intensity fluctuations of the beam. The photorefractive self-focusing effects used here have been described in detail in conjunction with photorefractive screening solitons^{14–18}. The formation of a bright screening soliton may be viewed in the following manner. A narrow light beam propagates in the centre of a biased dielectric (photorefractive) medium. In the illuminated region, electrons are optically excited, and therefore the conductivity increases and the resistivity decreases. Thus, the voltage drop occurs primarily in the dark regions leading to a large space-charge field $E_{sc}(\mathbf{r})$ there, $\mathbf{r}$ being the coordinates in the plane transverse to the propagation direction of the beam. The change in the refractive index $\Delta n(\mathbf{r})$ is linearly proportional to the space-charge field via the electro-optic (Pockels') effect. When the polarity of the field is properly chosen, $\Delta n(\mathbf{r})$ is proportional to $-E_{sc}(\mathbf{r})$ and this creates a 'graded index waveguide' that guides the beam that generated it¹⁶.

We have recently reported the first (to our knowledge) experimental observation of self-trapping of a partially spatially incoherent optical beam¹³; we used a quasi-monochromatic laser beam scattered off a rotating diffuser to generate a speckled beam¹⁹ in which the phases of any two well-separated points varied randomly with time. This speckled beam was self-trapped with diffraction being eliminated; the beam envelope maintained a constant diameter throughout propagation¹³. However, that previous experiment served only as a 'proof of principle' as the rotating diffuser generated partially spatially incoherent light only. The light origi-

nated from a laser and was temporally coherent (quasi-monochromatic) at all times.

Here we report self-trapping of a broad-band spatially and temporally incoherent light beam, emerging from an incandescent source; a light bulb. We have used a quartz-tungsten-halogen incandescent bulb to generate the white-light beam. The light is initially sent through a spectral filter to limit the frequency band to 380–720 nm (the temporal coherence time of the beam is of the order of a few femtoseconds). The light was collimated into a beam, sent through a polarizer to keep one polarization only, and focused onto the input face of an SBN:75 photorefractive crystal ($\text{Sr}_{0.75}\text{Ba}_{0.25}\text{Nb}_2\text{O}_6$). The spectrum of the light incident on the crystal is shown in Fig. 1. Our SBN crystal exhibited photorefractivity to wavelengths between roughly 380 and 520 nm, which gives a normalized bandwidth $\Delta\nu/\nu_0$ of roughly 0.3. In the crystal, the incoherent beam propagates along the crystalline *a*-axis with its polarization parallel to the *c*-axis (extraordinary polarization). We used a lens to image the beam at the input and output faces of the crystal onto a CCD (charge-coupled device) camera. As with photorefractive screening solitons¹⁸, the magnitude of the nonlinearity is fine-tuned with a uniform background beam by generating a bias level of electrons in the conduction band. For this purpose, we use an ordinarily polarized 488-nm laser beam which was expanded to illuminate the crystal uniformly. Self-focusing occurs with the application to the crystal of an appropriate voltage (magnitude and polarity) which gives rise to a space-charge field that has a large component along the *c*-axis, thus using the $r_{33} = 1,022 \text{ pm V}^{-1}$ electro-optic coefficient to create the index change required for self-trapping. An input beam of $14 \mu\text{m}$ (all beam sizes are given here as full-width at half-maximum, FWHM) diffracts to $82 \mu\text{m}$ after 6 mm of propagation. The large diffraction angle demonstrates the spatial incoherence of the beam. A coherent beam of size $14 \mu\text{m}$ at 380 nm wavelength would have diffracted to $35.34 \mu\text{m}$, whereas the same-size input beam at 720 nm would have diffracted to $63.1 \mu\text{m}$. Applying 600 V between the electrodes separated by 6 mm results in self-trapping of the beam, which traps to $12 \mu\text{m}$. The total optical power in the white-light input beam is 70.8 nW. As this power is spread over 340 nm, the amount of light at the photorefractively sensitive wavelengths is $\sim 25 \text{ nW}$. The 488-nm background beam had a power of 400 nW. These very low power levels, when translated to the corresponding beam intensities, result in a very long formation time (which is related to the dielectric relaxation time) for the self-trapped beam.

Our experimental results are shown in Figs 2 and 3. Figure 2a shows the profile of a $14 \mu\text{m}$ beam at the input face of the crystal. Figure 2b shows the profile of the $82 \mu\text{m}$ -wide normally diffracting beam in the absence of nonlinearity (zero voltage). Figure 2c–l shows the temporal evolution of the beam at the output face of the crystal that occurs once the nonlinearity (voltage) is turned on. The beam starts to self-focus by going through a quasi-steady-state regime that is reminiscent of quasi-steady-state photorefractive solitons¹⁰. Then, the beam breaks up and moves towards the positive *c*-axis and forms the steady-state self-trapped beam as shown by Fig. 2i–l. The centre of the self-trapped beam of Fig. 2j has moved a distance of $57 \mu\text{m}$ away from the centre of the initial diffracted beam, towards the *c*-axis. This 'displacement' of the self-trapped beam is closely related to the self-bending effects that all photorefractive solitons experience; it is driven by a diffusion field that is the lowest-order correction to the space-charge field that supports the solitons²⁰. The self-trapped beam roughly maintains its structure for at least eight hours, during which the beam fluctuated slightly in shape and drifted slowly towards the *c*-axis. It remained a self-trapped entity and did not break up or diminish for the entire duration of our experiment. Possible reasons for the slow fluctuations in the shape of the self-trapped beam are 'environmental' changes, such as temperature variations or slow drifting in the optical power emitted from the sources, as we have not used any

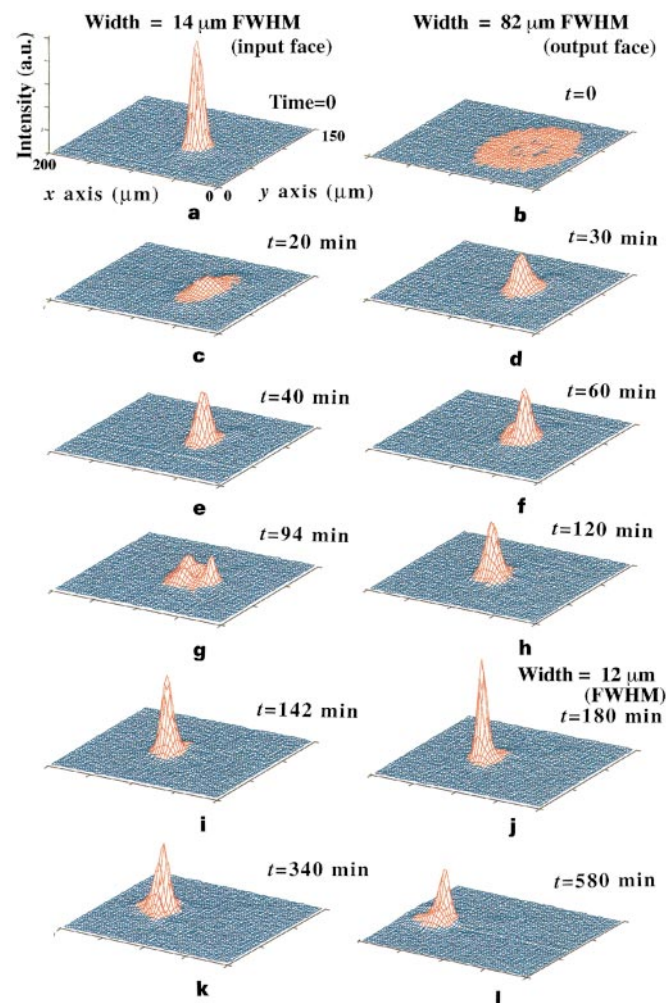


Figure 2 Steady-state self-trapping of a white-light beam. **a**, Profile of 14- μm beam (full-width at half-maximum, FWHM) at input face of crystal; **b**, profile of 82- μm beam at output face of crystal; **c–l**, temporal evolution of beam at output face of crystal.

special means to isolate our system. We emphasize, however, that the beam remained self-trapped (localized) for as long as we have monitored the experiment. We believe that the small fluctuations are not related to the self-trapping mechanism of the incoherent beam.

Quasi-steady-state trapping was observed even without use of background illumination. Figure 3a shows the profile of a 26- μm beam at the input of the crystal. Figure 3b shows the profile of the 100- μm normally diffracting beam in the absence of nonlinearity (zero voltage). Figure 3c–f shows the temporal evolution of the beam at the output face of the crystal. The beam focuses to a size of 26- μm then breaks up and eventually diminishes.

To our knowledge, this is the first observation of a self-trapped beam from a source that is both temporally and spatially incoherent. The phase across the self-trapped beam is varying in a random manner both in time and in space. Unlike the case of a self-trapped coherent beam, knowing (measuring) the phase at a particular point on the incoherent self-trapped beam cannot provide any phase information, even at very short distances away from that point. Furthermore, at each point on the beam, photons of widely varying frequencies coexist, so the absolute phase of the total optical field varies randomly between each two points separated by a distance of

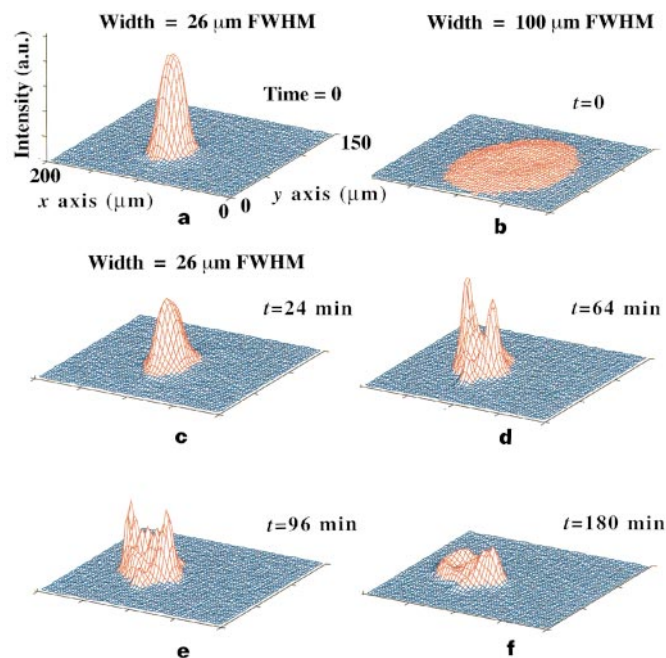


Figure 3 Quasi-steady-state self-trapping of a white-light beam. **a**, Profile of 26- μm beam at input face of crystal; **b**, profile of 100- μm diffracted beam at output face of crystal; **c–f**, temporal evolution of beam at output face of crystal.

the order of the optical wavelength. Yet, this incoherent white-light beam indeed self-traps.

Self-trapping of 'white' incoherent light introduces the possibility of using incoherent sources (for example, light-emitting diodes) for optical interconnects, beam steering, and other applications that have been thus far proposed only for (coherent) solitons. On the fundamental level, self-trapping of incoherent light raises many intriguing questions. We believe that the statistics of the self-trapped incoherent beam are affected by the self-trapping: going from the delocalized statistics of the input thermal source (that is, the coherence depends on coordinate difference only) into a state in which the statistics depend also on the absolute coordinate across the self-trapped beam. How are the brightness and entropy, which are related to the coherence properties of the beam, affected by self-trapping? If brightness is improved, it must come at the expense of energy loss, in the form of radiation or absorption. Are self-trapped incoherent beams able to maintain their identities as they undergo collisions with each other? These and other issues are currently under investigation. □

Received 4 February; accepted 6 May 1997.

- Chiao, R. Y., Garmire, E. & Townes, C. H. Self-trapping of optical beams. *Phys. Rev. Lett.* **13**, 479–482 (1964).
- Hasegawa, A. & Tappert, F. Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. *Appl. Phys. Lett.* **23**, 142–144 (1973).
- Mollenauer, L. F., Stolen, R. H. & Gordon, J. P. Experimental observation of picosecond pulse narrowing and solitons in optical fibers. *Phys. Rev. Lett.* **45**, 1095–1098 (1980).
- Snyder, A. W., Mitchell, D. J., Poladian, L. & Ladouceur, F. Self-induced optical fibers. *Opt. Lett.* **16**, 21–23 (1991).
- Zabusky, N. J. & Kruskal, M. D. Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.* **15**, 240–243 (1965).
- Bjorkholm, J. E. & Ashkin, A. Self-focusing and self-trapping of light in sodium vapor. *Phys. Rev. Lett.* **32**, 129–132 (1974).
- Barthelemy, A., Maneuf, S. & Froehly, C. Propagation soliton et auto-confinement de faisceaux laser par non linéarité optique de Kerr. *Opt. Commun.* **55**, 201–206 (1985).
- Aitchison, J. S. *et al.* Observation of spatial optical solitons in a nonlinear glass waveguide. *Opt. Lett.* **15**, 471–473 (1990).
- Segev, M., Crosignani, B., Yariv, A. & Fischer, B. Spatial solitons in photorefractive media. *Phys. Rev. Lett.* **68**, 923–926 (1992).
- Duree, G. *et al.* Observation of self-trapping of an optical beam due to the photorefractive effect. *Phys. Rev. Lett.* **71**, 533–536 (1993).

11. Torruellas, W. E. *et al.* Observation of two-dimensional spatial solitary waves in quadratic media. *Phys. Rev. Lett.* **74**, 5036–5039 (1995).
12. Tikhonenko, V., Christou, J. & Luther-Davies, B. Three dimensional bright spatial soliton collision and fusion in a saturable nonlinear medium. *Phys. Rev. Lett.* **76**, 2698–2701 (1996).
13. Mitchell, M., Chen, Z., Shih, M. & Segev, M. Self-trapping of partially spatially-incoherent light. *Phys. Rev. Lett.* **77**, 490–493 (1996).
14. Segev, M., Valley, G. C., Crosignani, B., DiPorto, P. & Yariv, A. Steady state spatial screening-solitons in photorefractive media with external applied field. *Phys. Rev. Lett.* **73**, 3211–3214 (1994).
15. Christodoulides, D. N. & Carvalho, M. I. Bright, dark and gray spatial soliton states in photorefractive media. *J. Opt. Soc. Am. B* **12**, 1628–1633 (1995).
16. Segev, M., Shih, M. & Valley, G. C. Photorefractive screening solitons of low and high intensity. *J. Opt. Soc. Am. B* **13**, 706–718 (1996).
17. Iturbe-Castillo, M. D., Marquez-Aguilar, P. A., Sanchez-Mondragon, J. J., Stepanov, S. & Vysloukh, V. Spatial solitons in photorefractive $\text{Bi}_{12}\text{TiO}_{20}$ with drift mechanism of nonlinearity. *Appl. Phys. Lett.* **64**, 408–410 (1994).
18. Shih, M. *et al.* Observation of two-dimensional steady-state photorefractive screening-solitons. *Electron. Lett.* **31**, 826–827 (1995); Two-dimensional steady-state photorefractive screening solitons. *Opt. Lett.* **21**, 324–326 (1996).
19. Goodman, J. *Statistical Optics* (Wiley, New York, 1985).
20. Singh, S. R. & Christodoulides, D. N. Evolution of spatial optical solitons in biased photorefractive media under steady state conditions. *Opt. Commun.* **118**, 569–576 (1995).

Acknowledgements. This work was supported by the US Army Research Office and the US National Science Foundation.

Correspondence should be addressed to M.S. (e-mail: segev@ee.princeton.edu).

Crystallization of hard-sphere colloids in microgravity

Jixiang Zhu*, Min Li†, R. Rogers‡, W. Meyer‡, R. H. Ottewill§, STS-73 Space Shuttle Crew||, W. B. Russel† & P. M. Chaikin*

* Department of Physics, † Department of Chemical Engineering, Princeton University, Princeton, New Jersey 08544, USA

‡ NASA Lewis Research Center, Cleveland, Ohio 44136-3191, USA

§ School of Chemistry, University of Bristol, Bristol BS8 1TS, UK

|| Johnson Space Flight Center, Houston, Texas 77058, USA

The structure of, and transitions between, liquids, crystals and glasses have commonly been studied with the hard-sphere model^{1–5}, in which the atoms are modelled as spheres that interact only through an infinite repulsion on contact. Suspensions of uniform colloidal polymer particles are good approximations to hard spheres^{6–11}, and so provide an experimental model system for investigating hard-sphere phases. They display a crystallization transition driven by entropy alone. Because the particles are much larger than atoms, and the crystals are weakly bound, gravity plays a significant role in the formation and structure of these colloidal crystals. Here we report the results of microgravity

experiments performed on the Space Shuttle Columbia to elucidate the effects of gravity on colloidal crystallization. Whereas in normal gravity colloidal crystals grown just above the volume fraction at melting show a mixture of random stacking of hexagonally close-packed planes (r.h.c.p.) and face-centred cubic (f.c.c.) packing if allowed time to settle^{7,8}, those in microgravity exhibit the r.h.c.p. structure alone, suggesting that the f.c.c. component may be induced by gravity-induced stresses. We also see dendritic growth instabilities that are not evident in normal gravity, presumably because they are disrupted by shear-induced stresses as the crystals settle under gravity. Finally, glassy samples at high volume fraction which fail to crystallize after more than a year on Earth crystallize fully in less than two weeks in microgravity. Clearly gravity masks or alters some of the intrinsic aspects of colloidal crystallization.

The thermodynamic phase diagram for hard spheres as obtained by computer simulations is: $\phi_{\text{liquid}} < \phi_{\text{freeze}} = 0.494 < \phi_{\text{coexist}} < \phi_{\text{melt}} = 0.545 < \phi_{\text{crystal}} < 0.74$ (refs 1–4), where ϕ is the volume fraction; the simulations also find reduced diffusion and a metastable glass phase for $\phi > 0.58$. For typical hard-sphere colloidal systems on Earth the gravitational length h is about 1–30 μm ($mgh = kT$, where m is the buoyant mass of a particle, g is the acceleration due to gravity, k is the Boltzmann's constant and T is temperature), below crystallite sizes, and much less than the sample sizes (of the order of centimetres). So it is not clear what effects sedimentation, concentration gradients and gravitational stresses have on the kinetics and thermodynamics of the crystals¹².

The particles consist of uniform poly(methylmethacrylate) (PMMA) spheres, 508 or 518 nm in diameter ($\approx 2a_0$), polydispersity $\sim 5\%$, with a thin (10-nm) grafted layer of poly(hydroxystearic acid) to prevent aggregation¹¹. The particles are suspended in an index of refraction (1.51) matching mixture of decalin and tetralin. Using a stirring bar in the sample cells, the samples were mixed (shear melted) by the astronauts on day 2 of the flight. The static laser light scattering (wavelength, $\lambda = 791 \text{ nm}$) was obtained with a 50- μm beam focused through a cylindrical sample and lens on a translucent screen and recorded with a video camera. The sample was translated to obtain data from many independent crystallites.

Crystal close packing can be obtained by stacking hexagonal planes of spheres. With a first layer as A, there are two equivalent placements B and C of the second layer above the interstitial sites of the first. If the stacking continues, the f.c.c. lattice is the arrangement ABCABC..., hexagonal close-packed (h.c.p.) is ABABAB..., but any arrangement has the same volume fraction, $\phi = 0.7404$. The random arrangement ABACBACBCA... with no repeating sequence

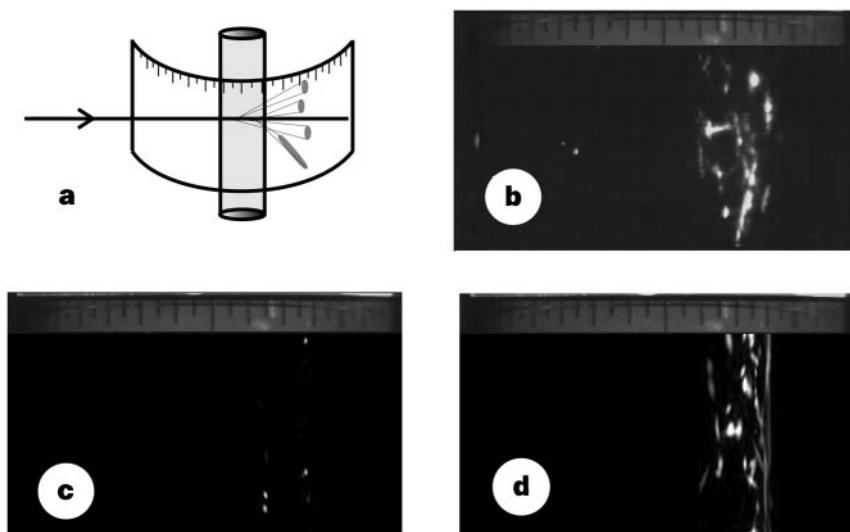


Figure 1 a, Bragg scattering geometry. A laser beam incident on a cylindrical sample is scattered onto a cylindrical screen. **b**, Bragg scattering from a sample with $2a_0 = 518 \text{ nm}$ and $\phi = 0.537$. Note the dominance of streaks rather than spots, indicating Bragg rods and a two-dimensional structure as might be expected for a random hexagonal close-packed (r.h.c.p.) structure. **c**, **d**, Computer-generated scattering pattern for f.c.c. (**c**) and r.h.c.p. (**d**) structures.

is r.h.c.p. Simulations indicate that the f.c.c. structure is most stable for ϕ near 0.7404. But just above melting, with $\phi \approx 0.545$, the situation is less clear^{13,14}. Experiments on hard-sphere colloidal crystals in normal gravity find a mixture of r.h.c.p./f.c.c. with more r.h.c.p.^{7,8}.

In Fig. 1b we show a typical CCD (charge-coupled device) image

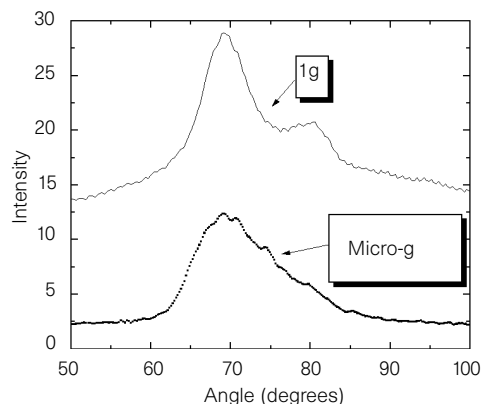


Figure 2 Scattering intensity as a function of angle obtained by averaging many scattering images as in Fig. 1b for a sample with $\phi = 0.537$. The peak in ground-based experiments at 80° is the (2,0,0) f.c.c. Bragg peak and shows that samples grown in normal gravity have some fraction of f.c.c. structure. Note the complete absence of any such f.c.c. structure in the microgravity data.



Figure 3 Left, photograph of a sample of 508-nm PMMA spheres in index-matching suspension at volume fraction $\phi = 0.504$, in the coexistence region of the phase diagram. In this pre-flight ground-based photograph the sample is phase-separated with 0.1-mm crystallites in the bottom half of the tube. Right, the same sample after recrystallization on the Space Shuttle in microgravity. Millimetre-sized crystals are dispersed in the liquid phase. Inset, higher-magnification view of the dendritic arms on several of the crystallites grown in microgravity.

recorded for the laser-Bragg screen (the scattering geometry is shown schematically in Fig. 1a). The unusual and characteristic feature of all of these diffraction patterns is the dominance of Bragg streaks (indicating two-dimensional order) rather than Bragg spots. The streaks result from incoherently stacked two-dimensional ordered sheets as one might expect for the r.h.c.p. structure. In Fig. 1c and d we show typical calculated images for f.c.c. and r.h.c.p. for our geometry and crystallite size¹⁵. Our data is clearly consistent with r.h.c.p. To test for the presence of f.c.c. we digitized the CCD images from many translations of the sample cell and computed the intensity I as a function of scattering wavenumber q (where $q = (2\pi/\lambda) \sin \theta/2$). In Fig. 2 we show $I(\theta)$ (unnormalized by the particle form factor) for the identical sample in the identical apparatus taken the same time after mix-melt in microgravity and normal gravity. The peak at 80° in the normal-gravity data is the (2,0,0) f.c.c. Bragg peak, completely absent in the microgravity data. Growth in microgravity yields almost pure r.h.c.p. crystals.

In Fig. 3 we show photographs of a sample in liquid–solid coexistence. On the left is the pre-flight photograph where the denser crystal phase settled below the liquid phase. On the right the same sample is shown after mix-melting and 3.6 days of crystallization in microgravity. The microgravity crystals are larger and exhibit dendritic arms (see inset) previously undetected and unpre-sumed. (Dendrites have been reported for charged colloids near a surface¹⁶.) The red and green colours indicate single crystals with different Bragg peaks which scatter visible white light with different colours in the same direction.

Should we expect dendrites? Ackerson and Schatzel^{17,18} explored the nucleation and growth of hard-sphere crystals experimentally and theoretically. Their model includes dynamics and kinetics

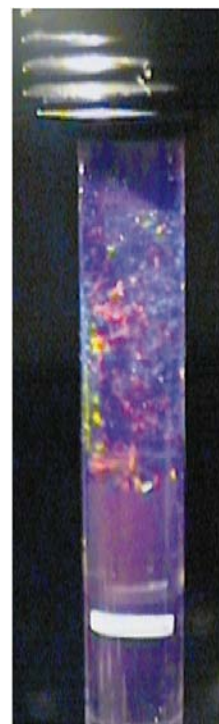


Figure 4 This sample, with $\phi = 0.619$ (deep in the 'glass' phase) completely crystallized in microgravity after failing to crystallize in normal gravity. The sample remained crystalline on return to ground. After 6 months, the bottom part of the sample was remelted using the stirring bar seen suspended in the glassy region. Two months later the bottom part of the sample remained glassy with neither nucleation nor growth in normal gravity.

particular to colloidal crystals. Our linear stability analysis of their equations¹⁹ indicates a dendritic instability at $\sim 13r_c$ (the critical nucleus $r_c \approx 3 \mu\text{m}$ for our system at $\phi = 0.504$) similar to the instability at $\sim 7r_c$ in molecular systems²⁰.

As a crystallite grows, its mass increases and in gravity it sinks faster. The viscous stresses applied to the crystallite (radius R) surfaces support its buoyant mass, M , and induce internal stresses of the order of $Mg/4\pi R^2$. When these stresses exceed the yield stress of the crystal, the crystal breaks. For a crystalline solid the yield stress, σ_{crit} is given approximately by βG , where G is the shear modulus and typically $\beta = 10^{-1}$ to 10^{-3} . Taking $M = (4\pi/3) R^3 \Delta\phi\rho$, $G \approx \chi kT/a_0^3$, $\Delta\phi = \phi_{\text{freeze}} - \phi_{\text{melt}} \approx 0.05$ and $a_0 = 0.25 \mu\text{m}$ we have $R_{\text{crit}} \approx 3\chi\beta kT/(a_0^3 \Delta\phi\rho) \approx 20h$, where h is the gravitational length ($\sim 30 \mu\text{m}$ for our samples) $\chi \approx 6$ (ref. 21), and ρ is the buoyant density of the particles. A sinking spherical particle can only grow to $\sim 600 \mu\text{m}$. Dendritic arms are sheared off much more readily. Sedimentation also sets up a flow field which can alter the diffusion field. For our samples, convection from sedimenting crystallites of radius R is faster than diffusion over the distance R when $R \geq$ several times a_0 , which may alter the growth mechanism.

The 'glass' sample that we studied had $2a_0 = 508\text{nm}$, $\phi = 0.619$, well into the glassy region observed in other ground-based systems^{22,23}. After mix-melting in normal gravity, the sample was allowed to sit undisturbed for a period of more than a year and never crystallized. The sample was mix-melted in microgravity and to our surprise began nucleation within 3.6 days. By the last experimental day of the shuttle flight the astronauts remarked on the centimetre sized crystals that had formed and filled the cell. This sample survived re-entry and sat in its fully crystalline state in our laboratory for 6 months. We then used the stirring bar in the sample cell to mix-melt the bottom half of the sample. For a month the crystalline region slowly grew into the disordered region and then essentially stopped. Figure 4 shows this sample two months after the bottom half was mix-melted. The glass region not in contact with the crystal has shown no evidence of nucleation. We therefore conclude that, at $\phi > 0.58$, the nucleation and growth of the crystals²⁴ are greatly hindered in a gravitational field. This may result from the geometry of random packings realized with and without applied stresses (for example, random close packing is $\phi = 0.64$, random loose packing, for systems under shear or gravity, is less well defined but is found to be $\phi \sim 0.58$; ref. 25).

We note that neither dendritic growth nor growth under sedimentation represent the slow quasi-equilibrium process that would assure an equilibrium crystal structure. The r.h.c.p. structure may be the equilibrium structure, but it may also result from non-equilibrium growth. The fact that f.c.c. packing was not found in microgravity but on crystalline samples brought back to normal gravity, or grown in normal gravity, suggests that stress may convert r.h.c.p. to f.c.c. One way to avoid these problems would be to use a temperature gradient to control the liquid-crystal interface (by an osmotic pressure gradient) and to 'zone refine' the samples in a microgravity environment. □

Received 7 January; accepted 6 May 1997.

1. Wood, W. W. & Jacobsen, J. D. Preliminary results from recalculation of the Monte Carlo equation of state of hard spheres. *J. Chem. Phys.* **27**, 1207–1208 (1957).
2. Adler, B. J. & Wainwright, T. E. Phase transition for a hard sphere system. *J. Chem. Phys.* **27**, 208–1209 (1957).
3. Hoover, W. G. & Ree, F. H. J. Melting transition and communal entropy for hard spheres. *Chem. Phys.* **49**, 3609–3617 (1968).
4. Alder, B. J., Hoover, W. G. & Young, D. A. Studies in molecular dynamic V. High density equation of state and entropy for hard disks and spheres. *J. Chem. Phys.* **49**, 3688–3696 (1968).
5. Woodcock, L. V. Glass transition and the hard sphere model. *J. Chem. Soc. Faraday II* **72**, 1667–1672 (1976).
6. Pusey, P. N. & van Meegen, W. Phase behavior of concentrated suspensions of nearly hard colloidal spheres. *Nature* **320**, 340–342 (1986).
7. Pusey, P. N., van Meegen, W., Bartlett, W. P. & Ackerson, B. J. Structure of crystals of hard colloidal spheres. *Phys. Rev. Lett.* **63**, 2753–2756 (1989).
8. Pusey, P. N. in *Liquids, Freezing and the Glass Transition* Course 10, *Colloidal Suspensions* (eds Hansen, J. P., Levesque, D. & Zinn-Justin, J.) 763–942 (Elsevier, Amsterdam, 1991).
9. de Kruijff, C. G., Jansen, J. W. & Vrij, A. in *Physics of Complex and Supramolecular Fluids* (eds Safran, S. A. & Clark, N. A.) 315–343 (Wiley-Interscience, New York, 1987).

10. Smits, C., Briels, W. J., Dhont, J. K. G. & Lekkerkerker, H. N. W. Influence of stabilizing coating on the rate of crystallization of colloidal systems. *Prog. Colloid Polym. Sci.* **78**, 287–292 (1989).
11. Phan, S. E. *et al.* Phase transition, equation of state, and limiting shear viscosities of hard sphere dispersions. *Phys. Rev. E* **54**, 6633–6645 (1996).
12. Biben, T., Ohnesorge, R. & Lowen, H. Crystallization in sedimentation profiles of hard spheres. *Europhys. Lett.* **28**, 665–670 (1994).
13. Frenkel, D. & Ladd, A. J. C. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres. *J. Chem. Phys.* **81**, 3188–3193 (1984).
14. Woodcock, L. V. Entropy difference between the face-centred cubic and hexagonal close-packed crystal structures. *Nature* **385**, 141–143 (1997).
15. Wilson, A. J. C. Imperfections in the structure of a cobalt II. Mathematical treatment of the proposed structure. *Proc. R. Soc. Lond. A* **180**, 277–285 (1942).
16. Gast, A. & Monovoukas, Y. A new growth instability in colloidal crystallization. *Nature* **351**, 553–557 (1991).
17. Schatzel, K. & Ackerson, B. J. Observation of density fluctuations during crystallization. *Phys. Rev. Lett.* **68**, 337–340 (1992).
18. Ackerson, B. J. & Schatzel, K. Classical growth of hard-sphere colloidal crystals. *Phys. Rev. B* **52**, 6448–6460 (1995).
19. Russel, W. B., Chaikin, P. M., Zhu, J., Meyer, W. V. & Rogers, R. Dendritic growth of hard sphere crystals. *Langmuir* (in the press).
20. Langer, J. S. Instabilities and pattern formation in crystal growth. *Rev. Mod. Phys.* **52**, 1–28 (1980).
21. Frenkel, D. & Ladd, A. J. C. Elastic constants of hard sphere crystals. *Phys. Rev. Lett.* **59**, 1169 (1987).
22. van Meegen, W. & Underwood, S. M. Glass transition in colloidal hard spheres: mode coupling theory analysis. *Phys. Rev. Lett.* **70**, 2766–2769 (1993).
23. van Meegen, W. & Underwood, S. M. Glass transition in colloidal hard spheres: measurement and mode coupling theory analysis of the coherent intermediate scattering functions. *Phys. Rev. E* **49**, 4206–4220 (1994).
24. Rintoul, M. D. & Torquato, S. Computer simulations of dense hard-sphere systems. *Chem. Phys.* **105**, 9258–9265 (1996).
25. Onoda, G. Y. & Liniger, E. G. Random loose packings of uniform spheres and the dilatancy onset. *Phys. Rev. Lett.* **64**, 2727–2730 (1990).

Acknowledgements. We thank B. J. Ackerson, P. N. Pusey and D. A. Huse for discussions. This work was supported by the NASA Microgravity Sciences program.

Correspondence and requests for materials should be addressed to P.M.C. (e-mail: chaikin@pupgg.princeton.edu).

Dissolved metals in surface sediment and a microbial mat at 100- μm resolution

W. Davison*, G. R. Fones* & G. W. Grime†

* Environmental Science Division, Institute of Environmental and Biological Sciences, Lancaster University, Lancaster LA1 4YQ, UK

† Microprobe Unit, Department of Nuclear Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

Sensors such as electrodes and optical fibre devices, optrodes, can be used to determine steep concentration gradients of chemical species in aquatic microenvironments, such as in the pore waters of surface sediments¹ and microbial mats^{2–4}, but are limited to a restricted range of determinands. The highest-resolution measurements of trace-metal concentrations in pore waters, at about 1.25 mm, have been provided by a recently developed thin-film gel technique^{5,6}, but the resultant metal distributions suggest that sub-millimetre-scale gradients need to be determined if the fluxes and cycling of the metals are to be fully quantified and understood. Here we report the development of this thin-film gel technique to measure Zn, Mn, Fe and As fluxes and concentrations at a resolution of 100 μm , and demonstrate the utility of the method *in situ* within the surface sediments and overlying microbial mat of a stream. Vertical profiles through the mat and sediments, and horizontal two-dimensional mapping just below the sediment–water interface, reveal the contrasting gradients, fluxes and remobilization niches of the four metal species at a sub-millimetre scale. The microbial mat appears to be an important regulator of the cycling of these metals. This technique has the potential to be extended to other chemical species and applied to other microenvironments with steep concentration gradients, such as redox boundaries, plant roots, animal burrows and sites of precipitation/dissolution in soils and sediments.

Microbial mats are laterally compressed ecosystems that support

most of the major biogeochemical cycles within a vertical dimension of only a few millimetres (ref. 7). The use of microsensors such as electrodes and optrodes has shown that there are steep concentration gradients (millimetres or less) of O_2 , $S(-II)$ and nitrogen species within microbial mats, reflecting the stratification of their communities^{2–4}. Although microbial mats are known to accumulate metals⁸, there have been no measurements of trace-metal concentrations within their pore waters.

In this study, we have used the technique of diffusive gradients in thin films (DGT). A layer of polyacrylamide hydrogel impregnated with Chelex chelating resin is overlain by first, a layer of hydrogel of known thickness (0.2–1 mm) and second, a cellulose nitrate membrane filter⁹. A plastic assembly ensures that only the filter contacts the solution. Metal ions freely diffuse through the 0.45- μm filter pores and the hydrogel (95% water); they are then bound by the chelating resin. After deployment the resin gel layer can be sliced, eluted with acid and metals determined by atomic absorption spectroscopy (AAS) or inductively coupled plasma mass spectrometry (ICP-MS). Deployment for a given time at a known temperature allows calculation of a simple arithmetic mean flux during deployment. In practice, apart from the first minute of deployment, this flux is unlikely to vary. If pore waters are effectively buffered by rapid resupply from a local source, such as a desorbing

solid phase, the measured flux can be quantitatively interpreted as a concentration.

With previous DGT assemblies spatial resolution was limited to 1 mm by the relatively large size of the resin beads, the difficulty of slicing a rubbery gel at submillimetre intervals, and the large size (total thickness 3–5 mm) of the assembly. We have overcome these difficulties by using a 200- μm -thick resin layer impregnated with 5% SPR-IDA (suspended particulate reagent-Iminodiacetate) chelating resin (Cetac Technologies Inc., Omaha, USA) with a 0.2- μm bead size. A 200- μm -thick layer of hydrogel is then sandwiched between the resin-gel layer and filter, and held between two 250- μm -thick Al_2O_3 ceramic plates, one of which has a 1×7.5 cm window open to the solution. The whole micro-DGT assembly resembles that of the prototype¹⁰, but has a total thickness of only 1 mm. After deployment, the resin gel is dried down using a conventional gel drier onto a cellulose nitrate filter. The filter becomes impregnated with the resin beads while maintaining the dimensional integrity of the deployed assembly. The masses of metal per unit area of the filter incorporating the dried gel were measured using proton induced X-ray emissions (PIXE)¹¹, using a similar procedure to the measurement of Fe and Mn in dried gels from diffusional equilibration in thin films (DET)¹². For measurements of vertical profiles, the 1- μm PIXE beam was rastered in

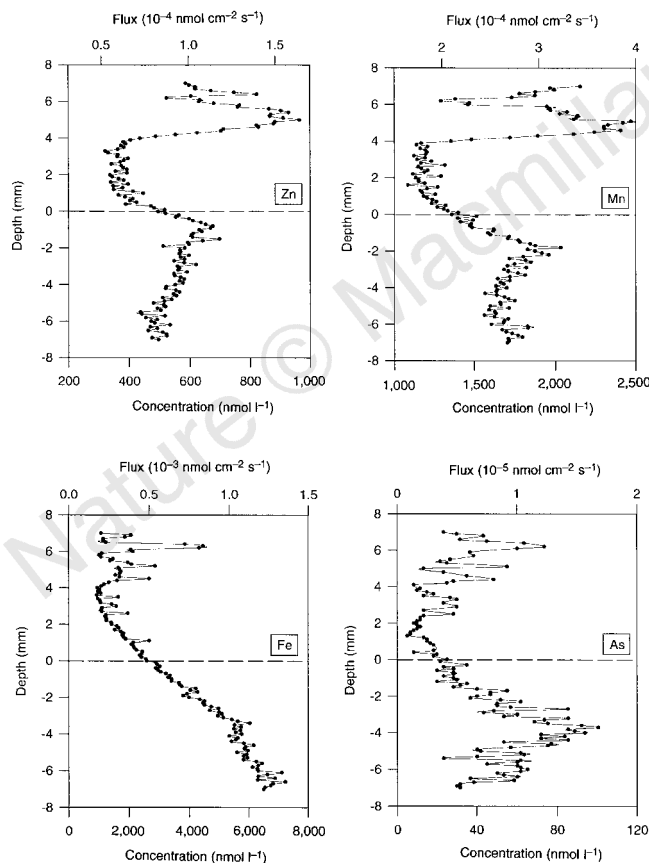


Figure 1 Concentrations of Zn, Mn, Fe and As, and fluxes of these elements to the DGT assembly, measured near the sediment surface in Sankey Brook on 5 June 1996 when the sediment was overlain by a microbial mat (5–6 mm thick). The position of the sediment surface (± 1 mm) is indicated by a broken line (at depth 0). Measurements were made by inserting a micro-DGT assembly vertically in the sediment. PIXE was used to measure the mass per unit area of metal in 100 (vertical) $\times$ 500 (horizontal) μm rasters. The PIXE calibration was confirmed to within 10% using DGT assemblies exposed to known concentrations of metals. The flux from the pore waters to the DGT assembly was measured directly. Pore-water concentrations are calculated assuming that they are well buffered *in situ*.

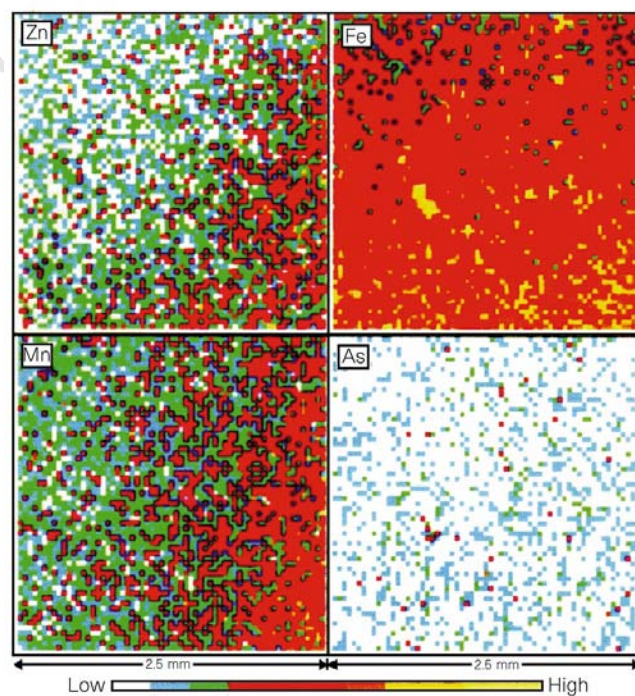


Figure 2 Two-dimensional representation of porewater relative concentrations of Zn, Mn, Fe and As in a plane perpendicular to the sediment surface at a position corresponding to the sediment–water interface. Concentration increases sequentially with the colour scale shown from white to yellow. The images were obtained by rastering a 1- μm proton beam (2 MeV) for 30 min in a 2.5 mm square on a micro-DGT assembly. The DGT assembly was deployed vertically in the surface sediment for 24 h.

100 × 500 μm rectangles adjacent to one another and average counts used for the rectangles. Two-dimensional image intensity maps of porewater concentrations were obtained by rastering the beam in 2.5-mm squares.

The micro-DGT assembly was deployed for 24 hours in a small polluted stream (Sankey Brook, Lancashire, UK; latitude 2° 30' W, longitude 53° 25' N). The sediment was overlain by a microbial community of *Oscillatoria*, presumably encouraged by shallow water (~60 cm), fairly stagnant conditions and a good nutrient supply. The location of the sediment–water interface and of the surface of the mat was assessed to within 1 mm, by the adherence of particles on the retrieved assembly. The resulting trace-metal profiles (Fig. 1) show clear evidence for remobilization of Zn and Mn near the surface of the microbial mat. The detailed similarity of the Zn and Mn profile strongly suggests that their mechanism of remobilization at the surface of the mat is related. By contrast the Zn and Mn maxima observed in the sediment were not coincident, indicating that their mechanisms of remobilization may be different at this location. The pronounced As maxima within the sediment is clearly separated from those of Zn and Mn. This ability of DGT to resolve remobilization sites for different metals should provide more detailed mechanistic information in future studies. There is some evidence for both Fe and As mobilization at the surface of the microbial mat, ~1 mm above the maxima for Zn and Mn.

At near neutral pH, Fe(II) and Mn(II) are the dominant redox states in solution which will be measured by DGT. The chelex resin binds cations and therefore will measure As(III) rather than As(V) which is present as oxyanions. Although As(OH)₃ is the dominant form in solution, its rapid dissociative equilibration with cationic species allows its measurement by DGT. Coincident As and Fe mobilization has been previously observed¹³. The detailed correspondence of As and Fe at submillimetre resolution within the microbial mat probably reflects reductive mobilization occurring within microniches. It is tempting to attribute the close correspondence of Zn and Mn to release of adsorbed Zn from reductively remobilized manganese oxyhydroxides, as bacteria can rapidly reduce Mn even in the presence of O₂ (ref. 14). However, the data provide no direct evidence of redox regulation and the behaviour of Zn in lakewater columns has been more closely linked to uptake by (and release from) biota than to iron and manganese transformations¹⁵. Alternative hypotheses could include metal release induced by pH changes associated with large and rapid changes in CO₂/O₂ as a result of photosynthesis and respiration.

The x-axes in Fig. 1 show fluxes and concentrations. DGT directly measures fluxes from the pore waters to the resin gel layer. Only if the pore waters are rapidly buffered can the results be interpreted as concentrations. Zn concentrations measured at the same site by conventional DGT at 1 mm intervals, using acid extraction, were shown to be independent of gel layer thickness. This result agrees with measurements in other systems, suggesting that Zn concentrations in pore waters are generally well buffered⁶. Without similar confirmation for other elements, their concentration scales must be treated cautiously. However, the steep concentration gradients observed for Mn and Zn can only be maintained if there is appreciable resupply from solid phase to solution, so for Mn at least, interpretation as a concentration is reasonable.

The image intensity maps (Fig. 2) were only obtained close to the sediment–water interface where concentration gradients are relatively uniform. They show clear horizontal gradients for both Mn and Zn, suggesting that the remobilization of these two elements in the sediment may be related even though their maxima are not coincident. Because of the horizontal heterogeneity, fluxes from the sediment to the overlying microbial mat calculated from concentration gradients will depend on the precise location of any measured concentration profile. There is no clear general relationship between As and Fe in this zone where there is a small concentration gradient and little evidence for active remobilization.

However, a region (~100 μm × 100 μm) with the highest iron measurements is in the same position as the zone of the highest As measurements, again indicating coincident mobilization of As and Fe within microniches.

Concentration gradients and deduced fluxes can only be quantitatively interpreted if measurements are made at sufficient spatial resolution which, in productive systems such as this, is 100 μm. Whether the deduced fluxes are generally representative of the sediment–water interface depend on the presence of horizontal concentration gradients and the temporal nature of the profile. The following geochemical interpretations apply only to the 24 hours during which the DGT measurement was made.

For iron there is no maximum within the surface 7 mm of sediment and a clear concentration gradient from pore waters to the overlying microbial mat is observed. There is similar transport of As, Zn and Mn across the sediment/microbial-mat boundary, but the location of the sub-surface supply is well defined. Clearly the lower part of the microbial mat acts as a sink for metal remobilized in the surface sediment and at the surface of the microbial mat. Microbial mats dominated by cyanobacteria have been shown to produce negatively charged polysaccharide biofloculants capable of binding trace metals¹⁶. It would be interesting to know if their production is restricted to the lower layers of the microbial mat.

Supply of fresh particulate material to the sediment surface is effectively cut off by the presence of the microbial mat. Measurements of Zn in Sankey Brook using 1-mm resolution DGT in March 1996, before a microbial mat was present, showed a much steeper and larger Zn maximum at the sediment–water interface. The mobilization of Zn and Mn at the microbial mat surface (Fig. 1) may simply be due to rapid remobilization of metals from decomposing particulate material landing on and being trapped by the microbial mat, although a more direct participation and control by the microbial community cannot be ruled out. The net effect of the microbial mat is therefore to limit the supply of metals to the sediment while simultaneously trapping any metal remobilized from the sediment. Consequently the amount of metal accumulated in the sediment will be less than it would have been in the absence of the microbial mat. According to these observations microbial mats may be active sites of metal remobilization as well as efficient traps for trace metals.

These measurements of trace metals at unprecedented high spatial resolution reveal geochemical processes occurring on a sub-millimetre scale. They indicate that sources and sinks of trace metals may be very localized, perhaps more so than physiologically driven oxygen gradients which have been elegantly measured by microsensors^{2,3}. Interpretation of metal fluxes and gradients and the role of microniches will be aided by the combined application of DGT and microsensors, particularly planar optrodes, which have provided fine-scale measurement of two-dimensional oxygen distributions⁴. Measurement of oxygen and labile concentrations of Fe(II), Mn(II) and S(–II) by voltammetric microelectrodes¹⁷ would complement the direct measurement by DGT of remobilization fluxes. We are at present developing DGT techniques for phosphate, caesium and sulphide. Combined applications of all these micro-measuring techniques are likely to advance our understanding of solute behaviour at any aquatic interface. □

Received 6 January; accepted 12 May 1997.

1. Sweets, J. P. R. A. & DeBeer, K. Microelectrode measurements of nitrate gradients in the littoral and profundal sediments of a meso-eutrophic lake (Lake Vechten, The Netherlands). *Appl. Environ. Microbiol.* **55**, 754–757 (1989).
2. Jørgensen, B. B. in *Microbial Mats* (eds Stal, L. J. & Caumette, P.) 243–254 (Springer, Berlin, 1994).
3. Glud, R. N., Jensen, K. & Revsbech, N. P. Diffusivity in surficial sediments and benthic mats determined by use of a combined N₂O–O₂ microsensor. *Geochim. Cosmochim. Acta* **59**, 231–237 (1995).
4. Glud, R. N., Ramsing, N. B., Gunderson, J. K. & Klimant, I. Planar optrodes: a new tool for fine scale measurements of two-dimensional O₂ distributions in benthic communities. *Mar. Ecol. Prog. Ser.* **140**, 217–226 (1996).
5. Davison, W. & Zhang, H. *In situ* speciation measurements of trace components in natural waters using thin-film gels. *Nature* **367**, 545–548 (1994).

6. Zhang, H. *et al.* In situ high resolution of fluxes of Ni, Cu, Fe, and Mn and concentrations of Zn and Cd in porewaters by DGT. *Geochim. Cosmochim. Acta* **59**, 4181–4192 (1995).
7. Paerl, H. W. & Pinckney, J. L. A mini-review of microbial consortia: Their roles in aquatic production and biogeochemical cycling. *Ecology* **31**, 225–242 (1996).
8. Xue, H.-B., Stumm, W. & Sigg, L. The binding of heavy metals to algal surfaces. *Wat. Res.* **22**, 917–926 (1988).
9. Zhang, H. & Davison, W. Performance characteristics of diffusion gradients in thin films for the in situ measurement of trace metals in aqueous solution. *Anal. Chem.* **67**, 3391–3400 (1995).
10. Davison, W., Zhang, H. & Grime, G. W. Performance characteristics of gel probe used for measuring pore waters. *Environ. Sci. Technol.* **28**, 1623–1632 (1994).
11. Davison, W., Grime, G. W. & Woof, C. Characterization of lacustrine iron sulfide particles with proton-induced X-ray emission. *Limnol. Oceanogr.* **37**, 1770–1777 (1992).
12. Davison, W., Grime, G. W., Morgan, J. W. & Clarke, K. Distribution of dissolved iron in sediment porewaters at submillimetre resolution. *Nature* **352**, 323–325 (1991).
13. DeVitre, R., Belzile, N. & Tessier, A. Speciation and adsorption of arsenic on diagenetic iron oxyhydroxides. *Limnol. Oceanogr.* **36**, 1480–1485 (1991).
14. Wehrli, B., Friedl, G. & Manceau, A. in *Aquatic Chemistry: Principles and Applications of Interfacial and Interspecies Interactions in Aquatic Systems* (eds Huang, P. *et al.*) 111–134 (Am. Chem. Soc., Washington DC, 1994).
15. Hamilton-Taylor, J. & Davison, W. in *Physics and Chemistry of Lakes* 2nd edn, Ch. 8 (eds Lerman, A., Imboden, D. & Gat, J.) 217–263 (Springer, Berlin, 1995).
16. Bender, J., Rodriguez-Eaton, S., Ekanemesang, U. M. & Phillips, P. Characterization of metal-binding biofloculants produced by the cyanobacterial component of mixed microbial mats. *Appl. Environ. Microbiol.* **60**, 2311–2315 (1994).
17. Brendel, P. J. & Luther, G. W. Development of a gold amalgam voltammetric microelectrode for the determination of dissolved Fe, Mn, O₂ & S(–II) in porewaters of marine and freshwater sediments. *Environ. Sci. Technol.* **29**, 751–761 (1995).

Acknowledgements. We thank H. Zhang for comments and A. Hunt for algal identification. This work was supported by the NERC.

Correspondence and requests for materials should be addressed to W.D. (e-mail: w.davison@lancaster.ac.uk).

Measuring the pulse of a plume with the sedimentary record

Nicky White* & Bryan Lovell††

* Bullard Laboratories, Madingley Rise, Madingley Road, Cambridge CB3 0EZ, UK

† BP Exploration, Britannic House, 1 Finsbury Circus, London EC2M 7BA, UK

Magmatic underplating associated with mantle plume activity is an important mechanism for driving regional surface uplift and denudation of large portions of the continents^{1,2}. Such uplift occurs rapidly because substantial volumes of basaltic melt are added to the crust over geologically short periods of time (1–10 Myr)², and can lead to large amounts of clastic sediment being shed into surrounding basins³. An intensively studied example of this process occurred in the North Sea basin during the Palaeogene period, where discrete pulses of deposition were triggered when sands were remobilized downslope from the shelf by turbidity currents and debris flows as a result of episodic changes of relative sea level³. Here we correlate the timing of these sediment pulses with the timing of surface uplift inferred to have been caused by episodic magmatic underplating on the continental shelf of northwestern Europe. This magmatism was related to activity of the Iceland plume, suggesting that individual pulses of sedimentation provide a potentially sensitive measure of plume activity, and so may be used to resolve time-dependent fluctuations in mantle plume activity predicted by theoretical studies of mantle convection.

We have chosen to examine the northwest shelf of Europe for three reasons. First, the history of vertical motions has been carefully studied³ over the past 20 years. Second, the distribution, composition, volume and age of onshore and offshore igneous activity are well documented^{4–7}. Third, the biostratigraphy, chronostratigraphy and seismic stratigraphy of Palaeogene sedimentary rocks have been systematically studied using both onshore data and extensive seismic and well data made available by the hydrocarbon industry^{3,8}.

In the Early Cenozoic era, an area encompassing the British Isles ($\sim 3 \times 10^5 \text{ km}^2$) underwent rapid exhumation (Fig. 1). Substantial quantities of Palaeogene muds and sands were deposited in the North Sea and in other basins including the Faeroe–Shetland basin, the Porcupine basin west of Ireland and southeastern England^{9–11}. Marine sedimentation was strongly pulsed, suggesting a pattern of transgression and regression which was first recognized by Lavoisier in 1766. A detailed history of highstands and lowstands of Palaeogene relative sea level in the North Sea basin is now available^{12,13}. These relative sea-level changes are recognizable around the British Isles but their magnitude and timing cannot be correlated throughout northwestern Europe or worldwide¹³. This observation indicates that global (that is, eustatic) sea-level variation did not control the cyclic pattern of deposition. We are especially concerned with the period between 62 and 54 Myr ago when a series of submarine fans formed at intervals of $\sim 1 \text{ Myr}$ on either side of the emergent Scottish landmass. These pulses of sand deposition are well documented in the North Sea and coeval pulses have been recognized in the Faeroe–Shetland basin¹⁴.

Permanent uplift and denudation of a region encompassing the British Isles were triggered by Palaeogene magmatic underplating^{1,15}. Geochemical and subsidence modelling show that this magma was generated by adiabatic decompression of asthenosphere with a potential temperature of 1,450–1,500 °C, requiring the existence of a mantle plume^{4,16}. The existence of a possible 50–250 m of transient dynamic uplift associated with the proto-Iceland plume^{2,17} cannot account for permanent uplift and does not affect our principal conclusions. Assuming Airy isostasy, the amount of magmatic underplating, X , required to generate an amount of denudation, D , is given by

$$X = \left(\frac{\rho_a - \rho_s}{\rho_a - \rho_x} \right) D + \left(\frac{\rho_a}{\rho_a - \rho_x} \right) T + \left(\frac{\rho_a - \rho_w}{\rho_a - \rho_x} \right) W \quad (1)$$

where T is the present-day topography (that is, residual uplift) and W is the water depth during the Late Cretaceous before uplift. Other parameters are given in Table 1. In the Irish Sea, $\sim 2.5 \text{ km}$ of strata have been eroded from extant Permian–Triassic basins and there is little residual uplift. If Late Cretaceous water depths were in the range 200–600 m, the required underplate is $8 \pm 2 \text{ km}$. If the same amount of underplating occurred beneath Scotland where Late Cretaceous water depths were 0–300 m, an average present-day topography of $\sim 500 \text{ m}$ implies that D is 0.5–2 km. This estimate of D is consistent with fission track analyses¹⁸ and with emplacement depths for igneous complexes⁴. If ρ_x is reduced to 2.8 Mg m^{-3} , then $X = 6 \pm 1 \text{ km}$.

Widespread magmatism is manifested at the surface by numerous central igneous complexes, lavas, linear dyke swarms, and other, minor, intrusive bodies (Fig. 1). Although accurate volumetric estimates of magmatism are difficult to make, studies of petrological fractionation support the notion that a substantial volume of magmatic underplating exists (3–5 times the volume inferred from outcrop¹). Large volumes are corroborated by forward and inverse modelling of concentrations of rare-earth elements for basaltic rocks. After corrections for fractional crystallization of the observed melt compositions, these data indicate that $\sim 5 \text{ km}$ of melt was produced at depth⁴. Wide-angle and deep seismic reflection data provide a further important constraint for the amount of underplating². The lower crust beneath the British Isles is strongly reflective and has a Poisson's ratio of ~ 1.86 , consistent with the existence of mafic sills. Beneath Scotland, the Lithosphere Seismic Profile in Britain (LISPB) wide-angle data set shows that P-wave velocity increases sharply at 20 km depth, and that the velocity 5–8 km above the Moho is $6.6\text{--}7 \text{ km s}^{-1}$ with calculated densities of $2.85\text{--}3.05 \text{ Mg m}^{-3}$ (ref. 19). Thus geochemical and geophysical observations are consistent with the underplating hypothesis.

Dating of onshore igneous rocks has proved difficult because of significant argon mobility and most of the published K–Ar ages are

† Present address: Bullard Laboratories, Madingley Rise, Madingley Road, Cambridge CB3 0EZ, UK.

uncertain. Reliable $^{40}\text{Ar}/^{39}\text{Ar}$ step-heating and Rb–Sr isochron ages^{5–7} have an analytical precision of 0.5–1.5 Myr but magnetic polarity data measurements can be used further to constrain the ages²⁰. Combined age and polarity data for all of the igneous centres show that magmatism lasted from 62 to 52 Myr ago, with a peak between 61 and 58 Myr.

The main pulses of submarine-fan sedimentation correlate with periods of most intense magmatic activity (Fig. 2). The phase of greatest fan development (Andrew) coincides with the acme of

magmatism between 61 and 58 Myr, within chron C26r. We infer that each phase of magmatism caused rapid uplift. Fan volumes and denudation estimates imply average erosion rates of $200 \pm 100 \text{ mm kyr}^{-1}$ during the period of interest, compatible with the range of global geomorphic estimates for fluvial systems²¹. More importantly, phased uplift gave rise to episodic lowstands of relative sea level. Each lowstand triggered the transport of sediment already on the shelf, downslope into deeper waters. The development of submarine fans in such a setting is particularly sensitive to change.

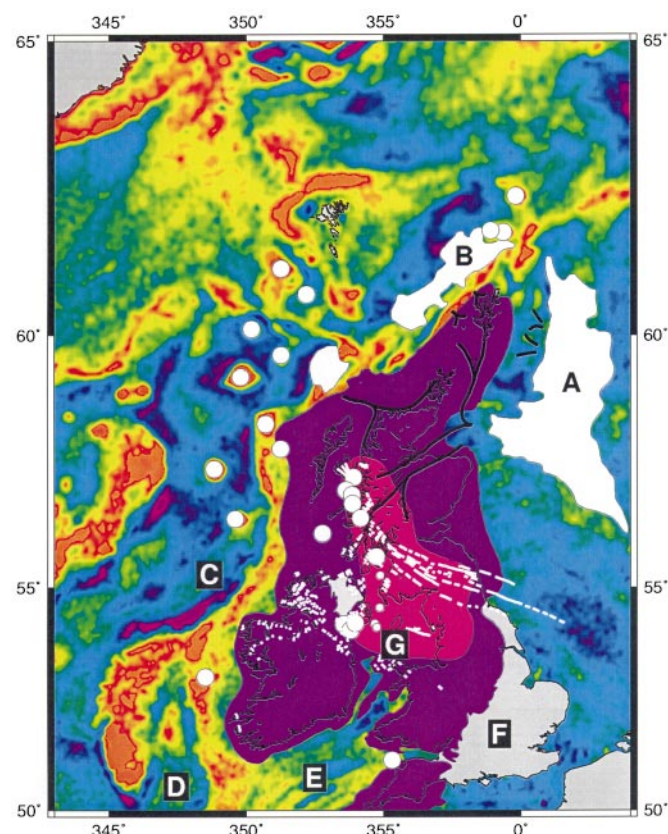


Figure 1 Satellite free-air gravity map of the British Isles (see Fig. 3 legend). The purple area encompassing the British Isles shows the region that underwent 0.5–1 km of denudation during the Palaeogene ($\sim 3 \times 10^5 \text{ km}^2$); the pink area shown underwent 2–3 km of denudation; the light grey patches offshore indicate submarine fan deposition in the North Sea and in the Faeroe–Shetland basin at time of maximum input (59 Myr ago: labelled ‘Andrew’ in Fig. 2); large white circles, major intrusive centres; small white circles and lines, minor intrusive centres, dykes and sills. A, North Sea; B, Faeroe–Shetland basin; C, Slyne–Erri basin; D, Porcupine basin; E, Celtic Sea basin; F, London–Hampshire basin; G, Irish Sea. Thick black lines indicate drainage pattern inferred from location of major channels offshore, from provenance studies, and from geomorphology. The denudation pattern is constrained by modelling subsidence analyses, vitrinite reflectance profiles, fission track data and sonic velocity information (refs 15, 18, and E. Rowley and K. Gallagher, personal communication). These data show that 1–3 km of denudation occurred rapidly during a period when climate was warm and relatively stable compared with the succeeding Neogene⁸. Increasing amounts of denudation have occurred in a transect from the Hampshire basin towards the Irish Sea where 2–3 km of strata have been removed. A maximum of $\sim 1 \text{ km}$ has been removed from most of Scotland. The submarine fan system shown in the North Sea consists of a large ($\sim 60,000 \text{ km}^2$) wedge of sand up to 700 m thick. Coeval fan systems in the northern North Sea and in the Faeroe–Shetland basin are smaller. Correlatable mineralogical units can be identified within Palaeogene submarine fan sequences. Systematic changes in sediment provenance reflect variations in source area composition, from Mesozoic and Palaeozoic sandstones to Moine and Dalradian metamorphic basement²⁶.

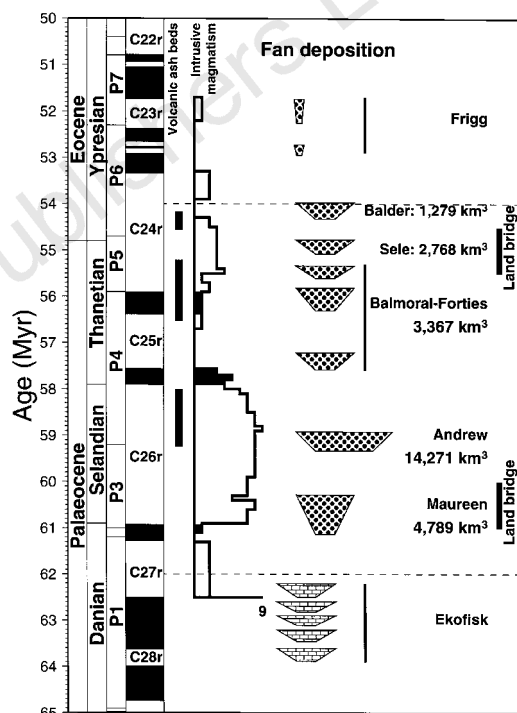


Figure 2 Stratigraphic chart based on timescale of Berggren *et al.*²⁰ and showing temporal relationship between intrusive igneous activity^{5–7}, offshore volcanic ash beds^{8,11}, and submarine fan deposition in North Sea¹³. At the left-hand side, stage names, planktonic foraminiferal biostratigraphic zonation and magnetostratigraphy are shown. The histogram of igneous activity (total of 35 ages) includes the total error estimate of age (monitor mineral is the Bern biotite 4B: $17.19 \pm 0.12 \text{ Myr}$). Magnetic polarity was used to position dates within the appropriate chron (black, normal polarity). Wedges with filled circles indicate aggregated submarine fan deposits. The nomenclature follows Neal¹³ and volume estimates are based on Reynolds²⁷. Horizontal dashed lines at 62 and 54 Myr mark initiation of the Iceland plume and the start of sea-floor spreading between Greenland and Europe, respectively. Note land bridges at start and end of 8-Myr period of interest (see Hooker in ref. 8). Danian rocks include submarine fans (wedges with brick shading which are labelled ‘Ekofisk’) which consist of reworked chalk derived from the erosion of Cretaceous strata that covered most of the British Isles³¹. The first major influx of clastic sediments (Maureen) occurred between 61.1 and 60.3 Myr ago. This fan system comprises $\sim 20\%$ of the total volume of submarine fan deposits in the Palaeogene. The succeeding Andrew fan system (59.4–58.9 Myr) represents $\sim 50\%$ of the total volume of submarine fan sediment with an implied minimum discharge of $1 \text{ m}^3 \text{ s}^{-1}$, equivalent to $7 \times 10^{10} \text{ kg yr}^{-1}$. Between 58 and 54 Myr, the Balmoral-Forties, Sele and Balder fan systems were generated with a combined volume of $\sim 20\%$. Total fan volume in the North Sea from 62 to 54 Myr is $\sim 26,000 \text{ km}^3$ ($\sim 6 \times 10^{16} \text{ kg}$)²⁷. By Early Ypresian times (54 Myr), submarine fan volumes were considerably smaller.

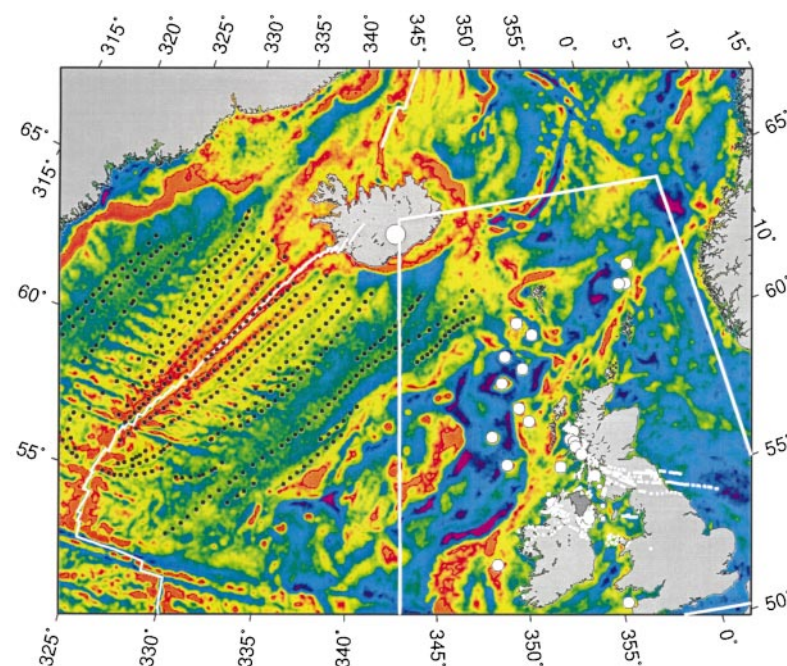


Figure 3 Satellite-derived free-air gravity map of Sandwell and Smith²⁸ for the North Atlantic Ocean (equal area projection). Red/yellow colours indicate gravity highs and purple/blue colours indicate gravity lows. The large white circle on Iceland indicates the location of the present-day plume centre at Vatnajökull (64.5° N, 17.3° W). Up to 500 km southwest along the Reykjanes ridge spreading centre, indicated by the solid white line, asthenospheric potential temperatures of $\sim 1,450^{\circ}\text{C}$ result in the generation of a 14-km-thick oceanic crust². Filled circles delineate prominent V-shaped ridges which transgress the magnetic anomaly pattern^{22,23}. These ridges are symmetrical about the spreading axis and converge southwards thus crossing progressively younger crustal isochrons. A less obvious set of V-shaped ridges occur north of Iceland. Isostatic calculations, wide-angle seismic data and rare-earth element modelling show that the crust is thicker beneath the most prominent ridges^{23,24}. V-shaped ridges are generated when pulses of anomalously hot asthenosphere flow down the spreading axis. The variation in crustal thickness indicates that asthenospheric potential temperatures beneath the plume fluctuate by $\pm 30^{\circ}\text{C}$ (ref. 24). Radial flow velocities required to produce V-shaped ridges can be calculated from Fig. 4. Features shown within white box are identified in Fig. 1.

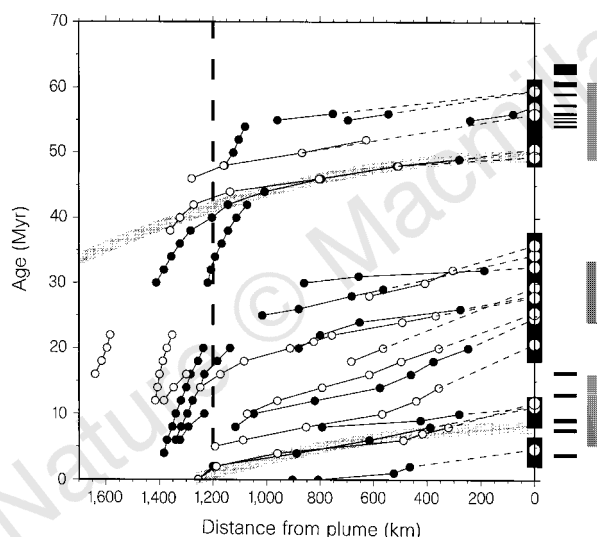


Figure 4 Age of V-shaped ridges plotted as a function of distance from the plume centre (updated from Vogt²²). At any given distance, the age of a ridge is determined from its intersection with the sea-floor spreading isochrons (for example, Müller *et al.*²⁹) which were calculated by linear interpolation between magnetic anomaly picks. Solid circles, ridges southeast of mid-ocean ridge; open circles, ridges northwest of mid-ocean ridge. Note approximate symmetry of ridges about spreading axis. Slope gives velocity of any high-temperature pulse as it travels down the spreading axis from northeast to southwest. These pulses travel ~ 800 km at an average velocity of $\sim 10\text{ cm yr}^{-1}$, with a rapid decrease to $\sim 1\text{ cm yr}^{-1}$ at $\sim 1,200$ km from the plume centre. Dashed lines and open circles extrapolate back to plume centre using a velocity which is averaged over last 3 points. Solid vertical bars at 0 km give range of starting times if velocity varies from 0 to 20 cm yr^{-1} . Solid horizontal bars (55–65 Myr), sediment pulses shown on Fig. 2; solid horizontal bars (0–20 Myr), radial plume pulsation from Wright and Millers³¹ seismic reflection and palaeoceanographic data analysis. Grey vertical bars, periods of minor upper-crustal shortening observed at continental margin north of Scotland³⁰; thick grey curves, calculated velocity of asthenospheric pulse assuming that plume flux is radial, varying according to equation (2) using parameters given in Table 1. At least six pulses have occurred during the past 30 Myr. V-shaped ridges appear to be absent between 35 and 50 Myr ago, indicating an hiatus in plume temperature fluctuation. Older diachronous ridges (40–60 Myr) are evident closer to the continental margins on either side of the ocean basin. Their velocities agree with those calculated for more prominent, younger, ridges. We suggest that at least three pulses originated at the plume centre between 60 and 50 Myr ago, taking ~ 1 Myr to travel for the plume centre to the British Isles, 500 km away.

in relative sea level^{8,9,12}. This well established sensitivity allows us to highlight the significance of a correlation between episodes of submarine fan deposition and uplift events probably caused by igneous activity. Modern and Quaternary studies demonstrate that the time lag between onset of a lowstand event and triggering of deposition is negligible for our purposes ($<10^5$ yr). The Palaeogene sedimentary record is proposed here as a measure of pulsed magmatic underplating, and can be regarded as a proxy for plume activity between 62 and 54 Myr ago. Once sea-floor spreading between Greenland and Europe began ~ 54 Myr ago, the continental margin moved away from the plume centre and submarine fan volumes in the North Sea decreased markedly^{3,9}.

The extent and magnitude of denudation suggest that the average

rate of melt addition beneath the entire British Isles was $0.1\text{ km}^3\text{ yr}^{-1}$. If the proportional change of submarine fan volume through time was triggered by surface uplift, which in turn was caused by variation in magmatic underplating, then the rate of melt addition fluctuated between 0.05 and $0.14\text{ km}^3\text{ yr}^{-1}$. This fluctuation could be produced by variation in asthenospheric potential temperature of $\pm 30^{\circ}\text{C}$. Our inference can be tested using independent but less well resolved evidence for fluctuations in plume activity over the past 60 Myr. Vogt^{22,23} first showed that prominent V-shaped ridges occur on the sea floor, south of Iceland (Fig. 3). Later work suggested that these ridges result from pulses of hotter asthenosphere travelling away from the centre of the plume²⁴. We determined the velocity and starting age of hot pulses of asthenosphere required to produce

Table 1 Parameters used in text

Symbol	Definition	Value
ρ_a	Density of asthenosphere	3.2 Mg m ⁻³
ρ_s	Density of sedimentary rock	2.6 Mg m ⁻³
ρ_w	Density of water	1 Mg m ⁻³
ρ_x	Density of magmatic underplating	2.9 Mg m ⁻³
$\dot{M}$	Mass flux of plume	2×10^{14} kg yr ⁻¹
h	Thickness of plume	100 km
R	Radius of plume	1,000 km

these ridges (Fig. 4). Downstream deceleration of each pulse is in accordance with the variation of residual depth and crustal thickness away from the plume centre²⁴.

For pulses travelling radially outwards from the plume centre, the distance of a pulse from the centre of the plume, R , at any time, t , is given by

$$R = \sqrt{\frac{\dot{M}t}{\pi h \rho_a}} \quad (2)$$

Using the parameter values given in Table 1, the geometry of the V-shaped ridges is consistent with rapid radial asthenospheric flow out to 1,200 km; there is not a requirement for pulses to be channelled exclusively along the Reykjanes ridge²⁵. This assertion is consistent with the geochemistry of ridge axis basalts if the radial flow beneath the ridge axis did not pass through the melting column beneath Iceland and if the plume centre is not located exactly on the ridge axis. Temperature variation within the core of the plume will also cause fluctuations in the pressure head at a mid-ocean ridge. Consequent changes in maximum compressive stress at the surrounding continental margins coincided with three discrete phases of mild folding and thrust faulting which occurred when pulsing of the plume was important (Fig. 4). However, these very modest amounts of crustal shortening cannot be responsible for Palaeogene uplift and denudation of the British Isles.

Several different data sets support our hypothesis that the existence of discrete pulses of Palaeogene submarine fan deposition is linked to mantle plume activity. Sedimentary deposits may therefore help to resolve otherwise inaccessible details of mantle plume activity on timescales of 0.5–1 Myr. We suggest that a global analysis of sedimentological data from hotspot settings would be fruitful. □

Received 5 December 1996; accepted 13 May 1997.

- Cox, K. G. Continental magmatic underplating. *Phil. Trans. R. Soc. Lond. A* **342**, 155–166 (1993).
- White, R. S. & McKenzie, D. Magmatism at rift zones: The generation of volcanic continental margins and flood basalts. *J. Geophys. Res.* **94**, 7685–7729 (1989).
- Parker, J. R. (ed.) *Petroleum Geology of Northwest Europe* (Geol. Soc., London, 1993).
- Brodie, J. A. Early Tertiary magmatism in the North Atlantic. Thesis, Univ. Cambridge (1995).
- Mussett, A. E., Dagley, P. & Skelhorn, R. R. *Time and Duration of Activity in the British Tertiary Igneous Province* 337–348 (Spec. Publ. 39, Geol. Soc., London, 1988).
- Mussett, A. E. & McCormack, A. G. *Magnetic Polarity Timescales: a New Test* 27–37 (Spec. Publ. 70, Geol. Soc., London, 1992).
- Pearson, D. G., Emeleus, C. H. & Kelly, S. P. Precise ⁴⁰Ar/³⁹Ar age for the initiation of Palaeogene volcanism in the Inner Hebrides and its regional significance. *J. Geol. Soc. Lond.* **153**, 815–818 (1996).
- Knox, R. W. O'B., Corfield, R. M. & Dunay, R. E. *Correlation of the Early Paleogene in Northwest Europe* (Spec. Publ. 101, Geol. Soc., London, 1996).
- Milton, N. J., Bertram, G. T. & Vann, I. R. *Early Palaeogene Tectonics and Sedimentation in the Central North Sea* 339–351 (Spec. Publ. 55, Geol. Soc., London, 1990).
- Shannon, P. M. *Early Tertiary Submarine Fan Deposits in the Porcupine Basin, Offshore Ireland* 351–373 (Spec. Publ. 62, Geol. Soc., London, 1992).
- Knox, R. W. O'B. *Tectonic Controls on Sequence Development in the Palaeocene and Earliest Eocene of Southeast England: Implications for North Sea Stratigraphy* 209–230 (Spec. Publ. 103, Geol. Soc., London, 1996).
- Stewart, I. J. in *Petroleum Geology of North West Europe* (eds Brookes, J. & Glennie, K.) (Graham & Trotman, London, 1987).
- Neal, J. E. in *Correlation of the Early Paleogene in Northwest Europe* (eds Knox, R. W. O'B., Corfield, R. M. & Dunay, R. E.) 15–42 (Spec. Publ. 101, Geol. Soc., London, 1996).
- Mitchell, S. M. et al. in *Petroleum Geology of Northwest Europe: Proc. 4th Conf.* (ed. Parker, J. R.) 1025–1034 (Geol. Soc., London, 1993).
- Brodie, J. & White, N. Sedimentary basin inversion caused by igneous underplating: Northwest European continental shelf. *Geology* **22**, 147–150 (1994).
- Scarrow, J. H. & Cox, K. G. Basalts generated by decompressive adiabatic melting of a mantle plume: a case study from the Isle of Skye, NW Scotland. *J. Petrol.* **36**, 3–22 (1995).

- Nadin, P. A., Kusznir, N. J. & Toth, J. Transient regional uplift in the early Tertiary of the northern North Sea and the development of the Iceland Plume. *J. Geol. Soc. Lond.* **152**, 953–958 (1995).
- Green, P. F., Duddy, I. R., Bray, R. J. & Lewis, C. L. E. in *Petroleum Geology of Northwest Europe: Proc. 4th Conf.* (ed. Parker, J. R.) 1067–1074 (Geol. Soc., London, 1993).
- Barton, P. J. LISBP revisited: a new look under the Caledonides of northern Britain. *Geophys. J. Int.* **110**, 371–391 (1992).
- Berggren, W. A., Kent, D. V., Aubry, M.-P. & Hardenbol, J. *Geochronology, Time Scales and Stratigraphic Correlation: Framework for an Historical Geology*. (Spec. Publ. 54, Soc. Econ. Geol. Palaeontol., Tulsa, 1995).
- Summerfield, M. A. & Hulton, N. J. Natural controls of fluvial denudation rates in major world drainage basins. *J. Geophys. Res.* **99**, 13871–13883 (1994).
- Vogt, P. R. Asthenosphere motion recorded by the ocean floor south of Iceland. *Earth Planet. Sci. Lett.* **13**, 153–160 (1971).
- Vogt, P. R. in *Structure and Development of the Greenland-Scotland Ridge* (eds Bott, M. H. P., Saxov, S., Talwani, M. & Thiede, J.) 191–213 (Plenum, 1983).
- White, R. S., Bown, J. W. & Smallwood, J. R. The temperature of the Iceland plume and origin of outward-propagating V-shaped ridges. *J. Geol. Soc. Lond.* **152**, 1039–1045 (1995).
- Ito, G., Lin, J. & Gable, C. W. Dynamics of mantle flow and melting at a ridge-centred hotspot: Iceland and the Mid-Atlantic Ridge. *Earth Planet. Sci. Lett.* **144**, 53–74 (1996).
- Morton, A. C., Hallsworth, C. R. & Wilkinson, G. C. in *Petroleum Geology of Northwest Europe: Proc. 4th Conf.* (ed. Parker, J. R.) 73–84 (Geol. Soc., London, 1993).
- Reynolds, T. Quantitative analysis of submarine fans in the Tertiary of the North Sea Basin. *Mar. Petrol. Geol.* **11**, 202–207 (1994).
- Sandwell, D. T. & Smith, W. H. F. Global marine gravity from ERS-1, Geosat and Seasat reveals new tectonic fabric. *Eos* **73**, 133 (1992).
- Müller, R. D., Roest, W. R., Royer, J.-Y., Gahagan, L. M. & Sclater, J. G. A Digital Age Map of the Ocean Floor. (Ref. Ser. 93–30, Scripps Inst. Oceanogr., La Jolla, 1993).
- Doré, A. G. & Lundin, E. R. Cenozoic compressional structures on the NE Atlantic margin: nature, origin and potential significance for hydrocarbon exploration. *Petrol. Geosci.* **2**, 299–311 (1996).
- Wright, J. D. & Miller, K. G. Control of North Atlantic Deep Water circulation by the Greenland-Scotland ridge. *Paleoceanography* **11**, 157–170 (1996).

Acknowledgements. We thank D. Aslanian, S. Bergman, J. Brodie, R. England, R. Knox, D. Lyness, S. Lewis, D. McKenzie, F. Nimmo, R. Rickaby and former colleagues of B.L. in BP Exploration, for their help. We also thank K. Cox and K. Gallagher for reviews. Figures were made using Generic Mapping Tools of P. Wessel and W. Smith. This work forms part of the PULSE (Paleogene Uplift Linked to Sediment Efflux) project funded by The British Petroleum Company plc and the Natural Environment Research Council.

Correspondence should be addressed to N.W. (e-mail: nwhite@esc.cam.ac.uk).

A new route for synthesis of dimethylsulphoniopropionate in marine algae

Douglas A. Gage*, **David Rhodes†**, **Kurt D. Nolte‡**, **Wayne A. Hicks***, **Thomas Leustek§**, **Arthur J. L. Cooper||** & **Andrew D. Hanson‡**

* Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824, USA

† Horticulture Department, Purdue University, West Lafayette, Indiana 47907, USA

‡ Horticultural Sciences Department, University of Florida, Gainesville, Florida 32611, USA

§ Center for Agricultural Molecular Biology, Rutgers University, New Brunswick, New Jersey 08903, USA

|| Department of Biochemistry, Cornell University Medical College, New York, New York 10021, USA

The 3-dimethylsulphoniopropionate (DMSP) produced by marine algae is the main biogenic precursor of atmospheric dimethylsulphide (DMS)^{1–3}. This biogenic DMS, formed by bacterial and algal degradation of DMSP^{4,5}, contributes about 1.5×10^{13} g of sulphur to the atmosphere annually³, and plays a major part in the global sulphur cycle, in cloud formation and potentially in climate regulation^{1,3}. Although DMSP biosynthesis has been partially elucidated in a higher plant^{6,7}, nothing is known about how algae make DMSP except that the whole molecule is derived from methionine^{8–12}. Here we use *in vivo* isotope labelling to demonstrate that DMSP synthesis in the green macroalga *Enteromorpha intestinalis* proceeds by a route entirely distinct from that in higher plants. From methionine, the steps are transamination, reduction and S-methylation to give the novel sulphonium compound 4-dimethylsulphonio-2-hydroxybutyrate (DMSHB), which is oxidatively decarboxylated to DMSP. The key

intermediate DMSHB was also identified in three diverse phytoplankton species, indicating that the same pathway operates in other algal classes that are important sources of DMS. The fact that a transamination initiates this pathway could help explain how algal DMSP (and thereby DMS) production is enhanced by nitrogen deficiency¹².

Many marine phytoplankton species and intertidal macroalgae accumulate DMSP as an osmolyte and cryoprotectant, particularly when salinity is high and nitrogen is limiting^{2,12,13}. We investigated DMSP synthesis first in the green macroalga *Enteromorpha intestinalis*, which is rich in DMSP ($\sim 20 \mu\text{mol g}^{-1}$ fresh weight) when grown in normal sea water. To identify biosynthetic intermediates between methionine (Met) and DMSP, we followed the metabolism of a small dose of ^{35}S -methionine. The main fates of ^{35}S Met were conversion to DMSP and incorporation into protein (Fig. 1a). Two stable compounds acquired ^{35}S rapidly and lost it as the ^{35}S Met dose was depleted, as expected for DMSP pathway intermediates; these were 4-methylthio-2-hydroxybutyrate (MTHB) and its S-methylated derivative DMSHB, a new natural product (Fig. 1b). The most plausible route from Met to these compounds is via the unstable 2-oxo acid, 4-methylthio-2-oxobutyrate (MTOB). We therefore tested for ^{35}S incorporation into MTOB by using a gentle extraction method (Table 1, method A) or by converting it to a stable derivative (Table 1, method B). Both approaches confirmed that MTOB acquired and lost ^{35}S in parallel with MTHB.

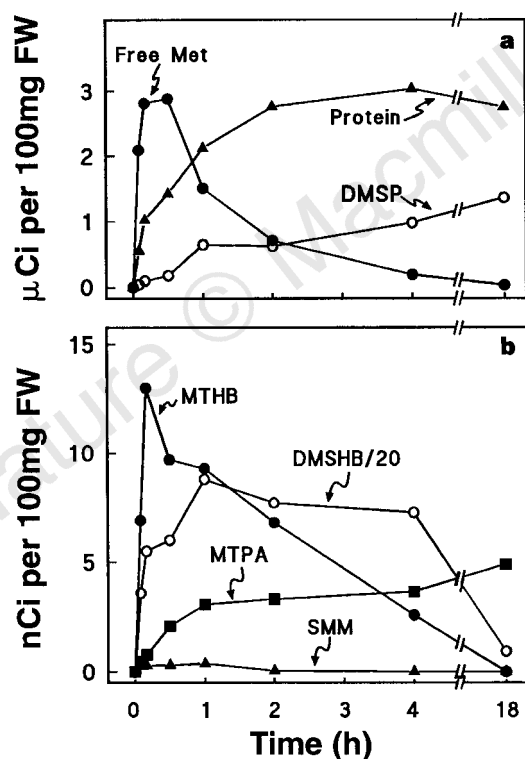


Figure 1 Labelling kinetics of selected small molecules and protein following administration of a tracer ^{35}S Met dose to *E. intestinalis*. Batches of fronds (100 mg fresh weight, FW) were supplied with $5 \mu\text{Ci}$ (0.5 nmol) ^{35}S Met. **a**, Free Met, DMSP and protein; **b**, Potential DMSP pathway intermediates; DMSHB data points are reduced by a factor of 20; MTPA, 3-methylthiopropylamine. No label (≤ 0.5 nCi per 100 mg) was detected in 3-dimethylsulphoniopropionaldehyde.

Computer modelling¹⁴ confirmed that the labelling kinetics of MTOB, MTHB and DMSHB were quantitatively consistent with roles as intermediates.

None of the other compounds monitored had labelling kinetics expected of intermediates in DMSP synthesis. S-Methylmethionine (SMM) and 3-dimethylsulphoniopropionaldehyde, both known to be DMSP synthesis intermediates in higher plants^{6,7}, acquired little or no ^{35}S (Fig. 1b). 3-Methylthiopropylamine, which has been proposed to be an intermediate in algae¹¹, labelled as would be expected for a minor end product (Fig. 1b). Labelling of another hypothetical intermediate, 3-methylthiopropionate^{8,11} varied from undetectable to comparable to that of MTOB in experiments with different lots of *E. intestinalis*. This can be attributed to a catabolic route involving oxidative decarboxylation of MTOB, which occurs in algae and other plants that do not contain DMSP, as well as in animals, and whose activity can vary with nutritional status^{9,15–17}.

These data support the pathway $\text{Met} \rightarrow \text{MTOB} \rightarrow \text{MTHB} \rightarrow \text{DMSHB} \rightarrow \text{DMSP}$, but are not consistent with a route involving SMM. To test this, ^{35}S -labelled SMM, MTHB and DMSHB were supplied (Table 2). SMM was scarcely metabolized, providing additional evidence that it is not an intermediate. MTHB was mainly converted to methionine, which in turn was incorporated into protein—a fate of MTHB well known in other plants and in animals^{18,19}. Some ^{35}S from MTHB also entered DMSP, consistent with MTHB being an intermediate; however, an indirect route via

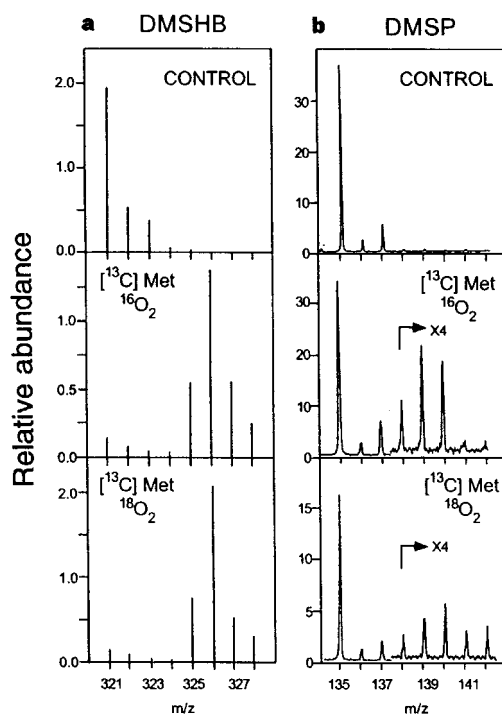


Figure 2 Evidence for conversion of DMSHB to DMSP by an oxygenase reaction in *E. intestinalis*. Fronds (100 mg) were incubated for 24 h with L-[U- $^{13}\text{C}_6$]Met (5 μmol) in flasks initially containing 21% $^{16}\text{O}_2$ or $^{18}\text{O}_2$ in N_2 ; controls received no Met and no $^{18}\text{O}_2$. **a**, SIM analysis of DMSHB. The peak areas in the control fit the expectation for natural-abundance C, S, Si and O isotopes. In the ^{13}C Met/ $^{16}\text{O}_2$ treatment, endogenous DMSHB was replaced by $^{13}\text{C}_4$ - and $^{13}\text{C}_6$ -labelled species (the S-demethylated derivatives of $^{13}\text{C}_6$ - and $^{13}\text{C}_8$ -DMSHB). Exposure to $^{18}\text{O}_2$ did not change this pattern, showing that the α -hydroxyl group does not originate from O_2 . **b**, FAB-MS analysis of DMSP. In the ^{13}C Met/ $^{16}\text{O}_2$ treatment, the peaks at m/z 139 and 140 are $^{13}\text{C}_4$ - and $^{13}\text{C}_6$ -DMSP. Under $^{18}\text{O}_2$, the signals at m/z 141 and 142 ($^{13}\text{C}_4$, ^{18}O - and $^{13}\text{C}_6$, ^{18}O -DMSP) show that 30–40% of the ^{13}C -labelled DMSP molecules contain an ^{18}O atom.

methionine cannot be excluded. As would be expected for a late intermediate in the pathway, DMSHB was metabolized efficiently to DMSP and to little else.

Stable isotope labelling was used to confirm and extend the radiotracer findings. We first verified the presence of the novel intermediate DMSHB and tested whether its conversion to DMSP involves an oxygenase reaction, perhaps analogous to that mediated by lactate oxidase²⁰. To do this, *E. intestinalis* was given [U-¹³C₅]Met in an atmosphere containing ¹⁶O₂ or ¹⁸O₂. Gas chromatography-mass spectrometry (GC-MS) with selected ion monitoring (SIM) revealed a small pool of DMSHB that became labelled with ¹³C in five or six positions, but not with ¹⁸O (Fig. 2a). Fast atom bombardment (FAB)-MS showed that most of the newly synthesized DMSP had four or five ¹³C atoms and that 30–40% of this DMSP became labelled with ¹⁸O (Fig. 2b). Note that recycling of the [¹³C₄]homocysteine moiety formed in methylation reactions would yield Met (and hence DMSHB and DMSP) with an unlabelled methyl group. The presence of such Met was confirmed by GC-MS, and modelling¹⁴ showed the observed proportions of [¹³C₄]- and [¹³C₅]DMSP conformed to expectations for recycling of the [¹³C₄]homocysteine moiety generated by the methylation step in DMSP synthesis. Also, photosynthetically generated ¹⁶O₂ lowers the intracellular abundance of ¹⁸O₂. The ¹³C and ¹⁸O data are thus fully consistent with a pathway in which the last step is an oxygenase-mediated oxidative decarboxylation of DMSHB.

Table 1 Metabolism of [³⁵S]Met by *E. intestinalis*

Extraction method	Incubation time (h)	³⁵ S incorporation (nCi per 100 mg)		
		MTOB	MTHB	DMSHB
A	0.5	17.3	6.8	51
	4.0	5.2	1.4	22
B	0.5	18.9	3.4	110
	4.0	≤2.0	0.6	35

Fronds (50 or 100 mg) were labelled and analysed as for Fig. 1, apart from the extraction (see Methods). The fronds extracted by methods A and B came from separate samples of algae.

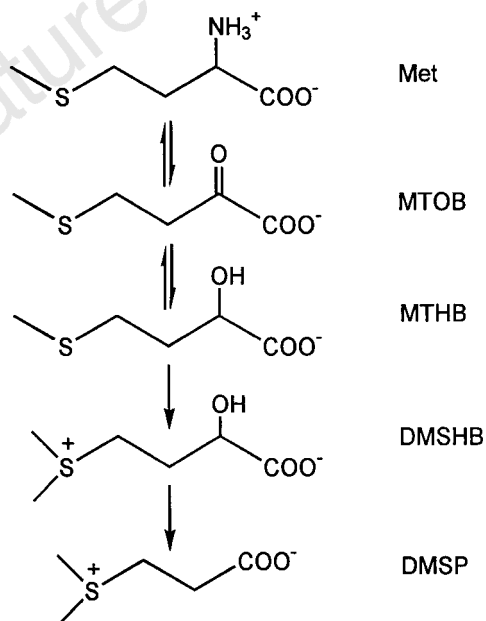


Figure 3 Proposed pathway of DMSP synthesis in *E. intestinalis* and other algae. The second step is shown as reversible because MTHB can be converted to Met (Table 2).

We also used stable isotope labelling to investigate the conversion of methionine to MTOB. *E. intestinalis* was given [¹⁵N]Met (5 μmol per 100 mg) and GC-MS was used to follow the labelling of amino acids. Glutamate acquired ¹⁵N readily (7.0% abundance at 2 h), as did aspartate and alanine, but the amide group of glutamine did not (<1% abundance at 2 h). This is consistent with transamination of methionine, but not with oxidative deamination: the ¹⁵NH₃ from a deamination reaction would have led to labelling of glutamine amide via the action of glutamine synthetase²¹. The operation of glutamine synthetase in *E. intestinalis* was confirmed by supplying ¹⁵NH₄⁺ (1 μmol per 100 mg) and showing that glutamine amide nitrogen was rapidly labelled (42% ¹⁵N abundance at 2 h). The conversion of MTHB to Met (Table 2) indicates that the Met → MTOB step is reversible, which is consistent with transamination²².

Collectively, the data for the chlorophyte macroalga *E. intestinalis* indicate that the pathway for synthesis of DMSP is that shown in Fig. 3. We tested for this pathway in marine planktonic species as these are major producers of DMSP and DMS^{1–3}. For this, we chose a prymnesiophyte (*Emiliania huxleyi*), a diatom (*Melosira nummuloides*) and a prasinophyte (*Tetraselmis* sp.). These algae all contained small pools of the key intermediate DMSHB which acquired label from [³⁵S]Met and lost it as the [³⁵S]Met was consumed (Fig. 4). All of them metabolized supplied [³⁵S]DMSHB to [³⁵S]DMSP (Fig. 4, inset). It is therefore likely that they have the same pathway as *E. intestinalis*.

Table 2 Uptake and metabolism of ³⁵S-precursors by *E. intestinalis*

Precursor	³⁵ S uptake (nCi)	³⁵ S incorporation (% of uptake)	
		DMSP	Protein Met
Met	454	6.2	58
SMM	451	≤0.2	≤1
MTHB	116	2.7	24
DMSHB	432	15.5	≤1

Data for Met are included for comparison. Fronds (100 mg) were incubated for 4 h with 500 nCi (1.1–1.4 nmol) of each compound. L-Isomers were supplied for all compounds, although these could have been racemized *in vivo*²⁰.

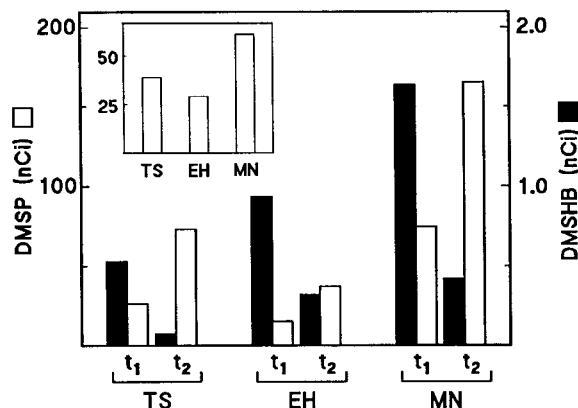


Figure 4 Evidence that DMSHB participates in DMSP synthesis in *Emiliania huxleyi* (EH), *Melosira nummuloides* (MN) and *Tetraselmis* sp. (TS). The main figure shows conversion of [³⁵S]Met to DMSP and DMSHB. Data are for 10⁷ (EH and TS) or 10⁵ cells (MN). Cultures were incubated with [³⁵S]Met (8–10 μCi; about 0.5 nmol) for a short time (t₁) to label heavily the pools of free Met and intermediates, and for a long time (t₂) to allow ³⁵S to chase from these pools. Times t₁ and t₂ (h), culture volumes (ml) and cell numbers were: EH, 1.5, 23, 12, 5 × 10⁷; MN, 2.5, 23, 1.2, 5 × 10⁵; TS, 0.25, 16, 25, 4.4 × 10⁷. The inset shows synthesis of [³⁵S]DMSP (nCi per 10⁷ cells) from [³⁵S]DMSHB (2–3 μCi, 0.2–0.3 nmol). Incubation was for 17–23 h. Culture volumes (ml) and cell numbers were: EH, 10, 5.5 × 10⁷; MN, 1.2, 5 × 10⁵; TS, 10, 1.4 × 10⁷.

Our data establish a pathway for DMSP biosynthesis in marine algae. This pathway has no steps in common with that in higher plants, which proceeds via SMM and 3-dimethylsulphonio-propionaldehyde^{6,7}. DMSP biosynthesis must therefore have evolved independently at least twice. Our results have two other implications. The first stems from the finding that a transaminase reaction stands at the head of the DMSP pathway; this may help explain why nitrogen deficiency enhances DMSP production^{12,23,30}. Depletion of cellular amino acids would favour the transamination reaction, thereby promoting DMSP synthesis when nitrogen is limiting. Second, our results suggest that DMSP may not be the only precursor of the DMS produced by living algae: DMSHB is another potential precursor *in vivo*. In support of this possibility, we have obtained preliminary evidence for extensive catabolism of supplied DMSHB to DMS in *Tetraselmis* sp. and *E. huxleyi*. □

Methods

Algae. *E. intestinalis* was collected in Florida and kept in aerated sea water at 18 °C in continuous fluorescent light (photosynthetic photon flux density, 50 µmol m⁻² s⁻¹). *M. nummuloides* (CCMP 482) and *Tetraselmis* sp. were cultured axenically at 25 °C in the above light regime in modified Gooday's medium²⁴ with 1 mM NO₃⁻; for *M. nummuloides*, 0.1 mM Na₂SiO₃ was added and Tris omitted. *E. huxleyi* (CCMP 373) was grown in f/2 medium²⁵ in daylight at 22 °C.

Radiochemicals. L-[³⁵S]SMM was synthesized enzymatically from L-[³⁵S]Met²⁶ or by treating L-[³⁵S]Met with 250 mM methanol in 6 M HCl at 110 °C for 4 h (ref. 27), and converted to L-[³⁵S]DMSHB with HNO₂ (ref. 28). [³⁵S]MTOB was made from L-[³⁵S]Met using L-amino acid oxidase, and converted to L-[³⁵S]MTHB using L-lactic dehydrogenase and NADH; [³⁵S]methylthiopropionate was obtained as a byproduct. Compounds were purified by ion exchange and TLE^{6,7}.

Labelling conditions. *E. intestinalis* fronds (50 or 100 mg) were incubated in 0.5–2.5 ml sterile sea water; ¹⁸O labelling was carried out in 50-ml flasks. For phytoplankton species, labelled compounds were added to growing cultures. Incubation was at 18–21 °C for *E. intestinalis* and *E. huxleyi* and 25 °C for *Tetraselmis* and *M. nummuloides*, under fluorescent light and with gentle agitation. Uptake of ³⁵S was estimated from its disappearance from the medium.

Metabolite analysis. Most metabolites were isolated by methanol–chloroform–water extraction, ion exchange, TLE and TLC^{6,7}. As MTOB broke down in these procedures, forming MTP, these compounds were isolated using 0.1 M HCl followed by ether extraction⁶ (method A) or by using paired samples extracted in 0.3 ml 66 mM NaBH₄ (to reduce MTOB to MTHB) or in pre-neutralized NaBH₄ (to give the endogenous MTHB level) (method B); MTOB was estimated by difference. Method B was also used to isolate 3-dimethylsulphonio-propionaldehyde as its hydroxy derivative⁷. Metabolites were identified by their respective lability and stability in cold 2 M NaOH. ³⁵S data were corrected for recovery, determined using labelled standards, and for products formed during extraction. Protein synthesis was estimated from ³⁵S labelling of the insoluble residue; the ³⁵S was shown to be in protein-bound Met by proteolysis and TLC.

Mass spectrometry. DMSP was analysed without derivatization by FAB-MS²⁹. DMSHB was derivatized as its *t*-butyldimethylsilyl ester/ether and analysed by GC-MS with SIM after on-column nucleophile-assisted S-demethylation²⁹. Authentic DMSHB⁷ was used to calibrate the SIM parameters. The diagnostic fragment ion cluster at *m/z* 321 (loss of a *t*-butyl radical) was monitored at the appropriate retention time. ¹⁵N-amino acids were analysed as described¹⁴, except that amide-¹⁵N abundance was determined using *N*-ethoxycarbonyl isobutyl esters by electron impact GC-MS³⁰.

Received 16 January; accepted 8 May 1997.

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* **326**, 655–661 (1987).
2. Keller, M. D., Bellows, W. K. & Guillard, R. R. L. in *Biogenic Sulfur in the Environment* (eds Saltzman, E. S. & Cooper, W. J.) 167–182 (American Chemical Society, Washington, DC, 1989).
3. Malin, G. in *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (eds Kiene, R. P., Visscher, P. T., Keller, M. D. & Kirst, G. O.) 177–189 (Plenum, New York, 1996).
4. Stefels, J. & van Boekel, W. H. M. Production of DMS from dissolved DMSP in axenic cultures of the marine phytoplankton species *Phaeocystis* sp. *Mar. Ecol. Progr. Ser.* **97**, 11–18 (1993).

5. Ledyard, K. M. & Dacey, J. W. H. Dimethylsulfide production from dimethylsulfoniopropionate by a marine bacterium. *Mar. Ecol. Progr. Ser.* **110**, 95–103 (1994).
6. Hanson, A. D., Rivoal, J., Paquet, L. & Gage, D. A. Biosynthesis of 3-dimethylsulfoniopropionate in *Wollastonia biflora* (L.) DC. Evidence that S-methylmethionine is an intermediate. *Plant Physiol.* **105**, 103–110 (1994).
7. James, F., Paquet, L., Sparace, S. A., Gage, D. A. & Hanson, A. D. Evidence implicating dimethylsulfoniopropionaldehyde as an intermediate in dimethylsulfoniopropionate biosynthesis. *Plant Physiol.* **108**, 1439–1448 (1995).
8. Greene, R. C. Biosynthesis of dimethyl-β-propiothetin. *J. Biol. Chem.* **237**, 2251–2254 (1962).
9. Pokorny, M., Marcenko, E. & Keglavic, D. Comparative studies of L- and D-methionine metabolism in lower and higher plants. *Phytochemistry* **9**, 2175–2188 (1970).
10. Chillemi, R., Patti, A., Morrone, R., Piatelli, M. & Sciuto, S. The role of methylsulfonium compounds in the biosynthesis of N-methylated metabolites in *Chondria coerulescens*. *J. Nat. Prod.* **53**, 87–93 (1990).
11. Uchida, A., Ooguri, T., Ishida, T., Kitaguchi, H. & Ishida, Y. in *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (eds Kiene, R. P., Visscher, P. T., Keller, M. D. & Kirst, G. O.) 97–107 (Plenum, New York, 1996).
12. Kirst, G. O. in *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (eds Kiene, R. P., Visscher, P. T., Keller, M. D. & Kirst, G. O.) 121–129 (Plenum, New York, 1996).
13. Blunden, G. & Gordon, S. M. Betaines and their sulphonio analogues in marine algae. *Progr. Phycol. Res.* **4**, 39–80 (1986).
14. Mayer, R. R., Cherry, J. H. & Rhodes, D. Effects of heat shock on amino acid metabolism of cowpea cells. *Plant Physiol.* **94**, 796–810 (1990).
15. Walters, D. S. & Steffens, J. C. Branched chain amino acid metabolism in the biosynthesis of *Lycopodium pennellii* glucose esters. *Plant Physiol.* **93**, 1544–1551 (1990).
16. Livesey, G. Methionine degradation: anabolic and catabolic. *Trends Biochem. Sci.* **9**, 27–29 (1984).
17. Scislowski, W. D., Hokland, B. M., Davis-van Thienen, W. I. A., Bremer, J. & Davis, E. J. Methionine metabolism by rat muscle and other tissues. *Biochem. J.* **247**, 35–40 (1987).
18. Dibner, J. J. & Knight, C. D. Conversion of 2-hydroxy-4-(methylthio)butanoic acid to L-methionine in the chick: a stereospecific pathway. *J. Nutr.* **114**, 1716–1723 (1984).
19. Mizayaki, H. H. & Yang, S. F. Metabolism of 5-methylthioribose to methionine. *Plant Physiol.* **84**, 277–281 (1987).
20. Walsh, C. *Enzymatic Reaction Mechanisms* (Freeman, New York, 1979).
21. Milfin, B. J. & Lea, P. J. Amino acid metabolism. *Annu. Rev. Plant Physiol.* **28**, 299–329 (1977).
22. Christen, P. & Metzler, D. E. *Transaminases* (Wiley, New York, 1985).
23. Groene, T. Biogenic production and consumption of dimethylsulfide (DMS) and dimethylsulfoniopropionate (DMSP) in the marine epipelagic zone: a review. *J. Mar. Syst.* **6**, 191–209 (1995).
24. Gooday, G. W. A physiological comparison of the symbiotic alga *Platymonas convolutae* and its free-living relatives. *J. Mar. Biol. Assoc. UK* **50**, 199–208 (1970).
25. Guillard, R. R. L. & Ryther, J. H. Studies on marine planktonic diatoms I. *Cyclotella nana* Hustedt and *Detonula confervacea* (Cleve.) Gran. *Can. J. Microbiol.* **8**, 229–239 (1962).
26. James, F., Nolte, K. D. & Hanson, A. D. Purification and properties of S-adenosyl-L-methionine: methionine S-methyltransferase from *Wollastonia biflora* leaves. *J. Biol. Chem.* **270**, 22344–22350 (1995).
27. Lavine, T. E., Floyd, N. F. & Cammaroti, M. S. The formation of sulfonium salts from alcohols and methionine in sulfuric acid. *J. Biol. Chem.* **207**, 107–117 (1954).
28. Zappia, V., Zydek-Cwick, C. R. & Schlenk, F. The specificity of S-adenosylmethionine derivatives in methyl transfer reactions. *J. Biol. Chem.* **244**, 4499–4509 (1969).
29. Gage, D. A. & Hanson, A. D. in *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (eds Kiene, R. P., Visscher, P. T., Keller, M. D. & Kirst, G. O.) 29–44 (Plenum, New York, 1996).
30. Huang, Z.-H. et al. Characterization of N-ethoxycarbonyl ethyl esters of amino acids by mass spectrometry. *J. Chromatogr.* **635**, 271–281 (1993).

Acknowledgements. We thank J. Hellebust, O. Schofield and J. S. Davis for algal material. This work was supported by grants from the NSF and ONR to A.D.H. and D.A.G., from the ONR to D.R. and T.L., from the NIH to the MSU-NIH Mass Spectrometry Facility and by the C.V. Griffin Sr Foundation. (Florida Agricultural Experiment Station journal series no. R-05740.)

Correspondence and requests for materials should be addressed to A.D.H. (e-mail: adha@gnv.ifas.ufl.edu) or D.A.G. (e-mail: gage@pilot.msu.edu).

Grazing-activated chemical defence in a unicellular marine alga

Gordon V. Wolfe*, Michael Steinke† & Gunter O. Kirst†

* College of Oceanic and Atmospheric Sciences, Oceanography Administration Building 104, Oregon State University, Corvallis, Oregon 97331-5503, USA

† University of Bremen, Marine Botany FB2, PO Box 330440, 28334 Bremen, Germany

Marine plankton use a variety of defences against predators, some of which affect trophic structure and biogeochemistry¹. We have previously shown² that, during grazing by the protozoan *Oxyrrhis marina* on the alga *Emiliania huxleyi*, dimethylsulphonio-propionate (DMSP) from the prey is converted to dimethyl sulphide (DMS) when lysis of ingested prey cells initiates mixing of algal DMSP and the enzyme DMSP lyase. Such a mechanism is similar to macrophyte defence reactions^{3,4}. Here we show that this reaction deters protozoan herbivores, presumably through the

production of highly concentrated acrylate, which has antimicrobial activity⁵. Protozoan predators differ in their ability to ingest and survive on prey with high-activity DMSP lyase, but all grazers preferentially select strains with low enzyme activity when offered prey mixtures. This defence system involves investment in a chemical precursor, DMSP, which is not self-toxic and has other useful metabolic functions. We believe this is the first report of grazing-activated chemical defence in unicellular microorganisms.

We examined five axenic strains of the prymnesiophyte *E. huxleyi*, all containing 75–110 mM internal DMSP. All strains synthesized constitutive DMSP lyase, which cleaves DMSP to form DMS, acrylate and a proton⁶. However, DMS production was negligible until cells were lysed by mechanical or chemical means². These strains were of similar size and growth characteristics but exhibited a bimodal distribution of *in vitro* DMSP lyase activity: strains 373 and 379 had hundreds-fold greater *in vitro* activity per cell (high activity) than strains 370, 374 and L (low activity) (Table 1).

The gross difference in enzyme activity among strains with similar DMSP titres provides a natural test for the role of this reaction as a herbivory deterrent. We selected protozoan grazers because microzooplankton are now recognized to be the dominant herbivores in marine planktonic systems⁷, especially for prey <20 µm in diameter, such as *E. huxleyi*. Predators tested included the dinoflagellate *O. marina* and protozoa isolated from Oregon

coastal waters, including a urotrichous ciliate and several flagellates. We conducted three types of grazing experiments: (1) single-prey trials with different *E. huxleyi* strains to look for gross toxicity, differences in grazing rates, or differences in predator growth rates; (2) feeding selectivity trials with prey mixtures of different *E. huxleyi* strains; and (3) feeding selectivity trials with prey mixtures of single *E. huxleyi* strains and other non-DMSP-containing prey species.

Some predators, notably *O. marina* and one flagellate, could ingest and grow on all five *E. huxleyi* strains (Fig. 1a), and DMS production during grazing reflected prey-strain DMSP lyase activity and was a reliable tracer of ingestion. DMS was not produced until *E. huxleyi* cells were actually ingested², suggesting that acrylate formation occurs inside predator food vacuoles. Because grazing on the high-activity strains 373 and 379 converted roughly 60% of total prey DMSP to DMS, we estimated production of ~70 mM acrylate inside predator food vacuoles per ingested cell. Such concentrated acrylate may have antimicrobial activity^{5,8}, and multiple prey were often observed together within a vacuole. Despite this, *O. marina* showed few negative effects. Clearance and grazing rates on high-activity strains were sometimes lower than for low-activity strains², but this was variable and might have been influenced by predator preconditioning.

However, grazers that tolerated high-activity strains avoided them when offered a choice of prey. This was evident in mixed-strain

Table 1 Characteristics of *E. huxleyi* strains examined

Strain and origin*	Low activity			High activity	
	370 North Sea	374 Gulf of Maine	L English Channel	373 Sargasso Sea	379 English Channel
Growth characteristics†					
Growth rate (d ⁻¹)	0.80	0.82	0.82	0.72	0.67
Diameter (µm)	4.4	4.9	4.7	5.3	4.8
DMSP (mM)	113	74	75	107	110
<i>In vitro</i> DMSP lyase‡					
Activity (fmol DMS cell ⁻¹ min ⁻¹)	0.01	<0.01	0.03	12.5	6.1
NaCl requirement (mM)	1,000	—	1,000	—	—
pH optimum	>8	5	5–6	6	6

All strains were axenic as shown by DAPI staining and plating out on 1% peptone seawater-agar, and none lithified under high-nutrient culture conditions. *In vitro* enzyme preparations were similar to those described in ref. 2 but were run with saturating levels of substrate (M.S. *et al.*, in preparation).

* Strain designations are from Provasoli-Guillard National Center for the Cultivation of Marine Phytoplankton (West Boothbay Harbor, Maine, USA) except for strain L. Synonyms: 370 (451B:F451), 373 (BT-6:CSIRO-CS-57), 374 (89E) and 379 (92A:P-92A).

† Conditions: 15 °C, 80 µmol m⁻² s⁻¹, 18 h light: 6 h dark cycle, f/2 medium.

‡ Conditions: 20 mM DMSP, 30 °C, NaCl optimum, 160 mM citric acid/phosphate buffer, pH 6.

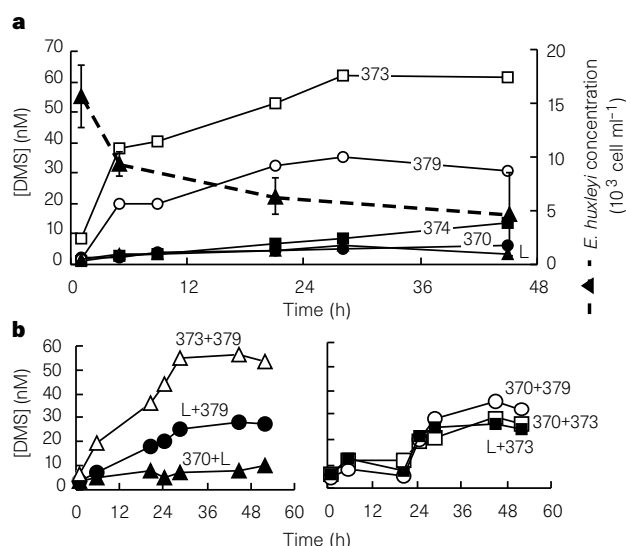


Figure 1 DMS production during grazing by *O. marina* on single (a) and multiple (b) *E. huxleyi* strains. Predators were maintained on the non-DMSP-producing chlorophyte *Dunaliella tertiolecta* and starved before the addition of *E. huxleyi* taken from exponentially growing batch cultures. Experiments were conducted in 250- or 500-ml polycarbonate bottles in 80 µmol m⁻² s⁻¹ at 15 °C (16 h light: 8 h dark cycle). DMS was sampled and analysed by gas chromatography, and cells were counted by epifluorescence microscopy². Single-strain initial prey concentrations were 15,000 ml⁻¹; grazing rates were similar on all five strains (mean and s.d. are shown). Multistrain initial prey concentrations were 20,000 ml⁻¹ (10,000 each strain) and 500 predators ml⁻¹. a, DMS production reflects *E. huxleyi* strain DMSP lyase activity (Table 1). b, Left, DMS production reflects grazing on pairs of high-activity strains (373 and 379) or low-activity strains (370 and L). b, Right, in three of four mixtures of high- and low-activity pairs, DMS production was low for the first 24 h, indicating selective feeding on the low-activity strain, then rose rapidly as predators switched over to the high-activity strain.

feeding trials in which we could not distinguish visually between *E. huxleyi* strains, but used DMS production to detect grazing of high-activity strains (Fig. 1b, right). Feeding discrimination against high-activity strains was also demonstrated in mixed-species trials, in which we found avoidance of high-activity *E. huxleyi* strains but ingestion of low-activity strains comparable to other prey species (Fig. 2). In some instances, non-DMSP-containing prey species were consumed preferentially and completely before high-activity *E. huxleyi* prey were grazed².

Other predators, including the urotichous ciliate and one gymnodinoid flagellate, grew on low-activity *E. huxleyi* strains, but could not subsist on either high-activity strain (data not shown). DMS production during grazing on high-activity strains was minimal, suggesting that there was little ingestion and lysis of prey. Short-term uptake experiments showed that, in the case of the urotichous ciliate, all *E. huxleyi* strains were captured readily by ciliate feeding currents, but prey seemed to be ingested selectively, with strong discrimination against high-activity strains (Table 2). Active postcapture prey rejection has been observed previously for tintinnid ciliates such as *Favella*^{9,10}.

Taken together these results suggest that the reaction may function as a deterrent against herbivory, and imply that some DMS production may be a by-product of algal chemical defence. A defensive role for acrylate produced by unicellular phytoplankton has been suggested previously¹¹, but we believe this to be the first evidence that it may achieve this through a grazing-activated mechanism. Such DMS production may be widespread among marine algal taxa: the release of DMS following macroalgal tissue damage has long been known⁶, and injury-activated discharge of

volatiles from the DMSP-synthesizing dinoflagellate *Amphidinium carterii* that might be involved in zooplankton deterrence has been observed¹².

Relatively little study has been given to specific mechanisms of chemical defence in planktonic microorganisms. Many bloom-forming phytoplankton taxa, especially cyanobacteria, prymnesiophytes and 'red-tide' dinoflagellates, are acutely toxic to meso- and microzooplankton^{13–15}, and ingestion of a few prey cells might suppress predator activities enough to confer a selective advantage to the clonal prey population. In contrast, the products of this reaction do not seem to be acutely toxic, at least to some grazers. However, *E. huxleyi* cells need not be fully ingested and lysed to deter predators, as high-activity strains were often avoided or rejected, in some cases shortly after capture (Table 2). Influencing predator selectivity may shift grazing pressure to other prey species and reduce competition for nutrients. It is well established that phagotrophic protozoa are highly discerning feeders, using chemosensory cues as well as prey morphology and motility to select¹⁶ or reject^{10,17} their prey. Such discrimination suggests that this reaction may play a signal role, possibly by generating DMS and acrylate gradients during prey handling which are sensed by predators, but which do not produce detectable bulk DMS. Preliminary video analysis suggests that *O. marina* reacts to gradients of acrylate with an increased rate of change in direction (R.C.D.I., deg s^{-1} ; 10^{-5} M acrylate R.C.D.I. = 279 ± 56 , $n = 10$; control R.C.D.I. = 181 ± 15 , $n = 80$; change in mean linear speed not significant; R. K. Zimmer-Faust, personal communication), consistent with a repulsion response¹⁸.

Plants that manufacture constitutive defence toxicants are at risk of self-toxicity, and also use metabolic energy producing compounds that may not be needed. Many macrophytes avoid these risks by storing toxicants in specialized organelles¹⁹ or by producing inducible toxins in response to predation²⁰. *E. huxleyi* avoids the self-toxicity of acrylate by maintaining it as DMSP, which is not only non-toxic²¹, but serves other useful cellular functions as an osmolyte²², a cryoprotectant²³ and a methyl donor²⁴. Cellular energy is therefore invested in a multi-use molecule that can be modified for instant defence whenever it is needed. Examples of marine defence compounds that serve additional ecological roles include diterpene alcohols, produced by the phaeophyte *Dictyota menstrualis*, which function as anti-fouling substances²⁵, and tetrodotoxin, produced by the puffer fish *Fugu niphobles*, which also serves as a sex pheromone²⁶. DMSP may be an example of a multifunctional defence precursor.

In addition to their roles in this proposed defence system, micromolar levels of DMS, DMSP and acrylate may act as chemo-attractants for marine bacteria^{27,28} and protists²⁹. These compounds may also function as chemical semaphore in trophic cascades: DMS emitted to the atmosphere seems to be a foraging cue for Antarctic procellariiform seabirds³⁰. We suggest that further roles

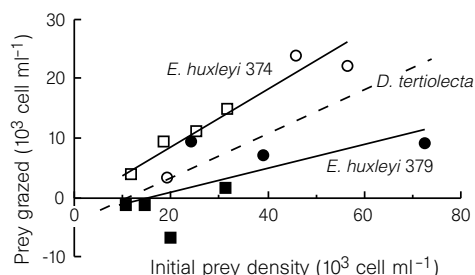
Table 2 Short-term uptake experiment

Treatment	Prey strain	<i>n</i> *	Ingested prey per predator
High-activity strains	373	22	1.8
	379	23	0.1
Negative control	fluorescent spheres	61	0.0
Low-activity strains	370	8	5.8
	374	11	3.1
	L	10	2.9
Positive control	<i>Chroomonas</i>	15	2.5

Results are for selective ingestion of low-activity *E. huxleyi* strains following capture by a urotichous ciliate. After 1-h incubations, preserved samples were stained with DAPI and filtered and the number of ingested prey cells per predator was observed by epifluorescence microscopy. In all instances, captured prey were seen attached to the exterior of ciliates (average, 2.9 prey per predator). Fluorescent spheres (PolySciences 3.46 μm fluoresbrite carboxylate) and the chlorophyte *Chroomonas* served as negative (captured but not ingested) and positive (captured and rapidly ingested) controls. Although captured by ciliates, high-activity *E. huxleyi* strains seemed to be ingested at much lower rates than low-activity strains, consistent with low DMS production during grazing and inability of ciliates to grow on high-activity strains as the sole food source.

* Number of predators examined.

Figure 2 *O. marina* feeding selectivity on mixtures of *Dunaliella tertiolecta* and either high- or low-activity *E. huxleyi* (strains 379 or 374), with prey ratios ranging from 20% to 80% *E. huxleyi*. Combined results from two experiments with initial total prey densities of 50,000 ml^{-1} (squares) or 100,00 ml^{-1} (circles) are shown; *O. marina* densities were 250–500 ml^{-1} . Bottles were incubated in the dark at 15 °C. At time zero and after 4 or 6 h, samples were fixed with alkaline Lugol's and ungrazed prey were counted by haemocytometer; grazed prey were calculated by difference and are plotted against initial prey concentrations. High-activity strain 379 (filled symbols) was grazed at much lower rates than low-activity strain 374 (open symbols), and within each experiment, uptake rates of strain 379 did not increase with increasing prey density (filled symbols). Grazing of *Dunaliella* (broken line; data points omitted for clarity) was consistent across treatments and between experiments, suggesting that presence of high-activity *E. huxleyi* does not reduce ingestion of other prey.



for these compounds as signals in marine trophic interactions remain to be discovered. □

Methods

Response of *O. marina* to acrylate gradients. *O. marina* cell suspensions were prepared on a microscope slide without a coverslip. Neutralized acrylic acid (Aldrich) was injected by micropipette to generate spatial gradients (equilibrium concentration, 10^{-5} M), and the cell response was observed for 30 s by phase-contrast microscopy and recorded on videotape using a NEC TI-23A CCD camera²⁸. A focal plane several millimetres from the surface of the slide was chosen to minimize wall effects. Cell swimming behaviour was analysed by a computer-assisted video motion analysis system (Motion Analysis model VP 110 with ExpertVision software), and rate of change in direction (R.C.D.I., deg s^{-1}), an analogue for angular velocity (turning frequency), and linear speed ($\mu\text{m s}^{-1}$) were calculated for 10–80 individual digitized cell paths. Control treatments were prepared with additions of seawater.

Received 17 January; accepted 10 April 1997.

- Verity, P. G. & Smetacek, V. Organism life cycles, predation, and the structure of marine pelagic ecosystems. *Mar. Ecol. Prog. Ser.* **130**, 277–293 (1996).
- Wolfe, G. V. & Steinke, M. Grazing-activated production of dimethyl sulfide (DMS) by two clones of *Emiliania huxleyi*. *Limnol. Oceanogr.* **41**, 1151–1160 (1996).
- Chew, F. S. in *Biologically Active Natural Products* (ed. Cutler, H. G.) 155–181 (American Chemical Society, Washington DC, 1988).
- Paul, V. J. & van Alstyne, K. L. Activation of chemical defenses in the tropical green algae *Halimeda* spp. *J. Exp. Mar. Biol. Ecol.* **160**, 191–203 (1992).
- Sieburth, J. M. Acrylic acid, an "antibiotic" principle in *Phaeocystis* blooms in Antarctic waters. *Science* **132**, 676–677 (1960).
- Cantoni, G. L. & Anderson, D. G. Enzymatic cleavage of dimethylpropiothetin by *Polysiphonia lanosa*. *J. Biol. Chem.* **222**, 171–177 (1956).
- Sherr, E. B. & Sherr, B. F. Bacterivory and herbivory: key roles of phagotrophic protists in prelude food webs. *Microb. Ecol.* **27**, 223–235 (1994).
- Slezak, D. M., Puskarić, S. & Herndl, G. J. Potential role of acrylic acid in bacterioplankton communities in the sea. *Mar. Ecol. Prog. Ser.* **105**, 191–197 (1994).
- Stoecker, D. K., Gallagher, S. M., Langdon, C. J. & Davis, L. H. Particle capture by *Favella* sp. (Ciliata, Tintinnida). *J. Plankton Res.* **17**, 1105–1124 (1995).
- Taniguchi, A. & Takeda, Y. Feeding rate and behavior of the tintinnid ciliate *Favella taraikaensis* observed with a high speed VTR system. *Mar. Microb. Food Webs* **3**, 21–34 (1988).
- Keller, M. D., Bellows, W. K. & Guillard, R. R. L. in *Novel Phytoplankton Blooms* (eds Cooper, E. M., Brice, J. V. M. & Carpenter, E. J.) 101–115 (Springer, Berlin, 1989).
- Wangersky, P. J. & Guillard, R. R. L. Low molecular weight organic base from the dinoflagellate *Alexandrium carteri*. *Nature* **185**, 689–690 (1960).
- DeMott, W. R., Zhang, Q.-X. & Carmichael, W. W. Effects of toxic cyanobacteria and purified toxins on the survival and feeding of a copepod and three species of *Daphnia*. *Limnol. Oceanogr.* **36**, 1346–1357 (1991).
- Hansen, P. J., Cembella, A. D. & Moestrup, O. The marine dinoflagellate *Alexandrium ostenfeldii*: paralytic shellfish toxin concentration, composition, and toxicity to a tintinnid ciliate. *J. Phycol.* **28**, 597–603 (1992).
- Uye, S.-I. & Takamatsu, K. Feeding interactions between planktonic copepods and red-tide flagellates from Japanese coastal waters. *Mar. Ecol. Prog. Ser.* **59**, 97–107 (1990).
- Verity, P. G. Feeding in planktonic protozoans: evidence for non-random acquisition of prey. *J. Protozool.* **38**, 69–76 (1991).
- Flynn, K. J., Davidson, K. & Cunningham, A. Prey selection and rejection by a microflagellate: implications for the study and operation of microbial food webs. *J. Exp. Mar. Biol. Ecol.* **196**, 357–372 (1996).
- Van Houten, J., Hauser, D. C. R. & Levandowsky, M. in *Biochemistry and Physiology of Protozoa* (eds Levandowsky, M. & Hutner, S. H.) **4**, 67–124 (Academic, New York, 1981).
- Fowden, L. & Lea, P. J. in *Herbivores: Their Interaction with Secondary Plant Metabolites* (eds Rosenthal, G. A. & Janzen, D. H.) 135–160 (Academic, New York, 1979).
- Harvell, C. D. The ecology and evolution of inducible defenses. *Q. Rev. Biol.* **65**, 323–340 (1990).
- Gröne, T. & Kirst, G. O. Aspects of dimethylsulfoniopropionate effects on enzymes isolated from the marine phytoplankter *Tetraselmis subcordiformis* (Stein). *J. Plant Physiol.* **138**, 85–91 (1991).
- Dickson, D. M. J. & Kirst, G. O. Osmotic adjustment in marine eukaryotic algae: the role of inorganic ions, quaternary ammonium, tertiary sulfonium and carbohydrate solutes. *New Phytol.* **106**, 645–655 (1987).
- Kirst, G. O. et al. Dimethylsulfoniopropionate (DMSP) in ice-algae and its possible biological role. *Mar. Chem.* **35**, 381–388 (1991).
- Ishida, Y. & Kadota, H. Participation of dimethyl- β -propiothetin in transmethylation reaction in *Gyrodinium aureolum*. *Bull. Jpn. Soc. Sci. Fish.* **34**, 699–705 (1968).
- Schmitt, T. M., Hay, M. E. & Lindquist, N. Constraints on chemically mediated coevolution: multiple functions for seaweed secondary metabolites. *Ecology* **76**, 107–123 (1995).
- Matsumura, K. Tetradotoxin as a pheromone. *Nature* **378**, 563–564 (1995).
- Sjölund, R. D. & Mitchell, R. Chemotactic responses of *Vibrio alginolyticus* to algal extracellular products. *Can. J. Microb.* **25**, 964–967 (1979).
- Zimmer-Faust, R. K., de Souza, M. P. & Yoch, D. C. Bacterial chemotaxis and its potential role in marine dimethylsulfide production and biogeochemical sulfur cycling. *Limnol. Oceanogr.* **41**, 1330–1334 (1996).
- Hauser, D. C. R., Levandowsky, M., Hunter, S. H., Chunosoff, L. & Hollwitz, J. S. Chemosensory responses by the heterotrophic marine dinoflagellate *Cryptothecodinium cohnii*. *Microbiol. Ecol.* **1**, 246–254 (1975).
- Nevitt, G. A., Veit, R. R. & Kareiva, P. Dimethyl sulphide as a foraging cue for Antarctic Procellariiform seabirds. *Nature* **376**, 680–682 (1995).

Acknowledgements. We thank B. Palenik for *E. huxleyi* strain L; L. Fessenden for protozoan isolates; R. Zimmer-Faust for the video motion analysis of *O. marina*; and E. Sherr for comments on the manuscript. This work was supported by grants from NASA and the European Community.

Correspondence and requests for materials should be addressed to G. Wolfe (e-mail: wolfe@ucs.orst.edu).

Effects of sea-ice extent and krill or salp dominance on the Antarctic food web

V. Loeb*, V. Siegel†, O. Holm-Hansen‡, R. Hewitt§, W. Fraser||, W. Trivelpiece|| & S. Trivelpiece||

* Moss Landing Marine Laboratories, P.O. Box 450, Moss Landing, California 95039-450, USA

† Institut für Seefischerei, Bundesforschungsanstalt für Fischerei, Palmaille 9, 22767 Hamburg, Germany

‡ Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0202, USA

§ National Marine Fisheries Service, Southwest Fisheries Science Center, P.O. Box 271, La Jolla, California 92038, USA

|| Polar Oceans Research Group, Department of Biology, Montana State University, Bozeman, Montana 59717, USA

Krill (*Euphausia superba*) provide a direct link between primary producers and higher trophic levels in the Antarctic marine food web^{1–6}. The pelagic tunicate *Salpa thompsoni* can also be important during spring and summer through the formation of extensive and dense blooms^{6–9}. Although salps are not a major dietary item for Antarctic vertebrate predators^{7,10}, their blooms can affect adult krill reproduction and survival of krill larvae. Here we provide data from 1995 and 1996 that support hypothesized relationships between krill, salps and region-wide sea-ice conditions^{11,12}. We have assessed salp consumption as a proportion of net primary production, and found correlations between herbivore densities and integrated chlorophyll-*a* that indicate that there is a degree of competition between krill and salps. Our analysis of the relationship between annual sea-ice cover and a longer time series of air temperature measurements^{12,13} indicates a decreased frequency of winters with extensive sea-ice development over the last five decades. Our data suggest that decreased krill availability may affect the levels of their vertebrate predators. Regional warming and reduced krill abundance therefore affect the marine food web and krill resource management.

In the Elephant Island area near the Antarctic Peninsula (Fig. 1) both krill and salps exhibited large interannual abundance fluctuations between 1976 and 1996 (Table 1). In general, larger krill population densities were encountered during the earlier part of the data set; densities from 1984–85 until 1995–96 were on average an order of magnitude less than during previous years. Randomization tests on an analysis of variance¹⁴ indicate that differences in abundance between these two periods are statistically significant and support lower krill abundance in recent years. In contrast, the highest salp densities occurred during three summers within the 1984–96 period.

Interannual fluctuations in krill abundance result largely from variations of year-class success: the highest population densities (for example, 1981–82) result from extremely good recruitment from the previous spawning season¹¹. Relatively low densities in 1984–85, 1990–91 and 1994–95 followed two or three years of poor and intermediate krill recruitment. Good recruitment is positively correlated with early seasonal spawning (in December–February), and both are positively correlated with extensive sea-ice in the Antarctic Peninsula region the preceding winter ($n = 17$; Kendall's $T = 0.40$, $P < 0.05$; $n = 12$, $T = 0.48$, $P < 0.05$). Poor recruitment and late spawning (in March) are associated with reduced regional sea-ice formation.

In contrast to krill, salp abundance is negatively correlated with extensive sea-ice. Unlike krill, which have lifespans of more than 5 years¹⁵, *S. thompsoni* live less than one year, and their fluctuations in abundance reflect annual variability in conditions promoting

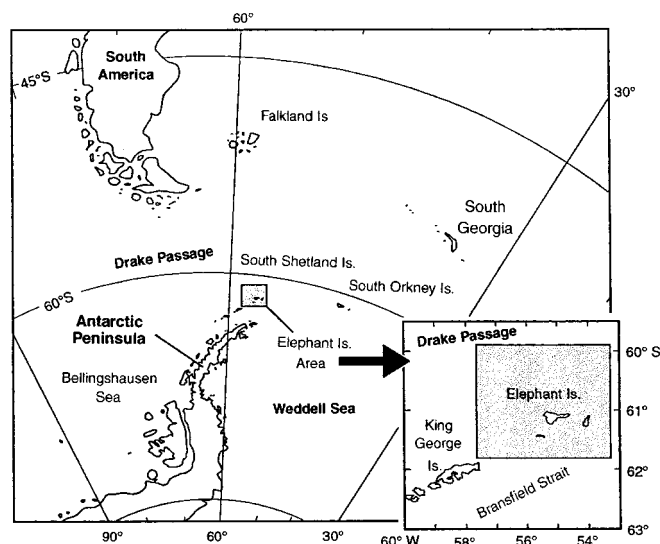


Figure 1 Location of the Elephant Island survey area near the Antarctic Peninsula. The major Antarctic krill harvesting activities occur within the southwest Atlantic sector between the Antarctic Peninsula and 30° W. This sector also supports large populations of resident and summer migrant krill-dependent species.

explosive population growth^{7,16}. Exceedingly high salp densities in 1983–84, 1988–89, 1989–90, 1992–93 and 1993–94 followed winters with relatively low sea-ice development. Mean salp abundance is negatively correlated with regional indices of sea-ice cover during the preceding winter ($n = 12$, $T = -0.58$, $P < 0.01$).

The differing relationships with sea-ice may be due to the ability of krill (late larvae, juveniles and adults) to exploit available ice algae during the winter months^{3,17}. In contrast, salps are obligate filter feeders¹⁶ that may benefit from open-water conditions that promote rapid population growth in early spring. We hypothesize that: (1) winters with extensive sea-ice cover inhibit spring salp blooms and promote early seasonal krill spawning; (2) these factors ensure good survival of the larval krill and subsequent recruitment to the population at the age of one year; and (3) a second winter of extensive ice cover amplifies the effect by promoting continued early spawning of krill and ensuring good survival of larvae spawned the preceding spring and summer^{11–13}. Conversely, we hypothesize that winters with poor sea-ice development promote extensive salp blooms and poor krill spawning. We also postulate that if the above effects are real, a long-term warming trend in the Antarctic Peninsula region could have profound effects on the dominance of krill and salp populations and, by implication, krill-dependent predators and the fate of newly fixed carbon.

Sea-ice coverage and duration in the Antarctic Peninsula region were above average during the winters of 1994 and 1995 (Table 1). As expected, salp densities were low the following summers. Krill density was low during summer 1994–95, and was dominated by large mature forms at least four years old, which is evidence of poor recruitment since the 1990–91 spawning season. However, unlike the previous three years, these krill had an early spawning season, which started in early December, and virtually all mature females were gravid or spent by mid-February 1995. Spawning success was indicated by the appearance of early-stage krill larvae in January and widespread occurrence of dense larval krill patches in February. High krill densities in 1995–96 arose mainly from the large numbers of one-year-old juveniles, and reflect exceedingly good overwintering survival and recruitment of the 1994–95 year class (Table 1). Early spawning and the presence of krill larvae during January 1996 followed above-average sea-ice conditions during winter 1995.

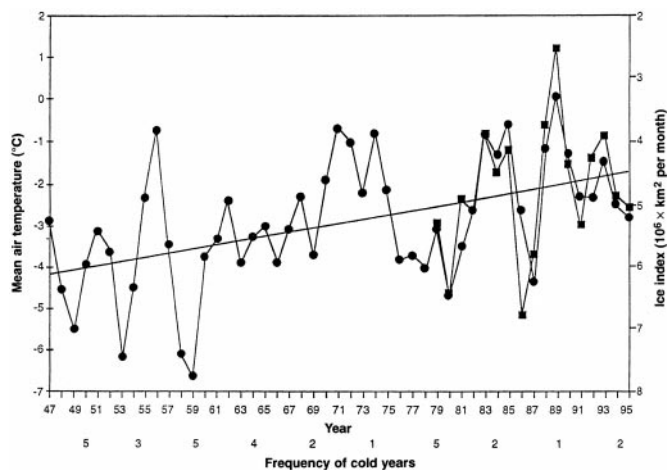


Figure 2 Long-term trends in mean annual surface air temperature in the Antarctic Peninsula region 1947–1995 (circles) and the regional sea-ice index 1979–1995 (squares). Frequency of cold years reflects the number of years per respective five-year period that mean air temperature was $\leq -2.5^\circ\text{C}$, the temperature associated with the median sea-ice index value during 1979–1995.

Krill and salp abundance relations during summer 1994–95 and 1995–96 differ greatly from those in 1993–94, when salps dominated the macrozooplankton (Table 2). During summer 1993–94, large concentrations of salps were encountered across the survey area, and median salp biomass (mg C m^{-2}) was an order of magnitude greater than that of krill. Using a daily consumption rate for salps of 25% of their body carbon¹⁸, the median uptake of daily primary production during January 1994 was 100 mg C m^{-2} , which was 19% of the average daily net primary production rate ($537 \text{ mg C m}^{-2} \text{ d}^{-1}$). This level of consumption is probably characteristic of other major 'salp years' such as 1989–90 and 1992–93 (Table 1). During the less dense blooms of 1983 and 1988, salp consumption of primary production was estimated at 10% and 9%, respectively^{18,9}. Consumption by salps during summer 1995 was negligible ($2\text{--}3 \text{ mg C m}^{-2} \text{ d}^{-1}$, compared with the mean primary production value of $1,318 \text{ mg C m}^{-2} \text{ d}^{-1}$) and, based on the low population density, consumption by krill was probably substantially lower than in the preceding years. The high chlorophyll-*a* values observed during 1995 (Table 1) may have resulted from low consumption by both krill and salps; conversely, low chlorophyll-*a* values during 1996 may have resulted from grazing by abundant krill. Significant negative correlations (Kendall's *T* tests, $P \leq 0.01$) between individual cruise chlorophyll-*a* concentrations and krill and salp densities during 1990–1995 ($n = 12$, $T = -0.58$, and $n = 8$, $T = -0.71$, respectively) suggest that grazing pressure by both krill and salps may affect phytoplankton standing stock size in the area. If this is true then much of the primary production during 1995 was left to be taken up by other herbivores (such as copepods) or was lost from the photic zone by flux into deeper waters¹⁶.

Although the estimated proportion of net primary production taken up by large salp densities during the peak summer phytoplankton bloom seems to be modest ($<20\%$), the impact of salp grazing during earlier, lower productivity spring bloom periods may be quite significant. Rapid spring population growth of *S. thompsoni* occurs in October and November; during years with extraordinarily large population growth, the grazing pressure by salps on relatively low spring phytoplankton standing stocks [primary production values $<50 \text{ mg C m}^{-2} \text{ d}^{-1}$ (ref. 19)] is no doubt quite significant. It has been suggested¹¹ that strong feeding competition by salps during

Table 1 Trends in krill and salp abundance

Field season	Krill				Salps		Chlorophyll- <i>a</i>		Sea-ice index (10 ⁶ km ² -months)
	<i>N</i>	Mean (No. 1,000 m ⁻³)	Spawning index	Recruitment index	<i>N</i>	Mean (No. 1,000 m ⁻³)	<i>N</i>	Mean (mg m ⁻²)	
1975-76	-	-	-	0.144	24	862.4	-	-	-
1976-77	-	-	-	0.048	-	-	-	-	-
1977-78	68	132.9	0.032	-	53	3.6	-	-	-
1978-79	-	-	-	0.069	-	-	-	-	-
1979-80	-	-	-	0.559	-	-	-	-	5.29
1980-81	33	49.5	0.896	0.757	42	24.1	-	-	6.43
1981-82	52	510.9	0.824	0.663	52	25.3	-	-	4.90
1982-83	12	90.6	-	0.119	-	-	-	-	5.12
1983-84	36	67.4	-	0.214	39	167.6	-	-	3.90
1984-85	103	12.3	0.748	0.175	115	42.1	-	-	4.48
1985-86	-	-	-	0.633	27	79.5	-	-	4.15
1986-87	-	-	-	0.291	-	-	-	-	6.77
1987-88	56	66.6	0.674	0.275	19	1.2	-	-	5.79
1988-89	50	41.7	0.144	0.063	22	200.0	-	-	3.76
1989-90	98	15.4	0.263	0.099	46	3,489.0	73	46.7	2.56
1990-91	81	4.8	0.802	0.587	-	-	85	65.2	4.35
1991-92	130	25.3	0.427	0.012	63	94.3	136	32.3	5.33
1992-93	136	26.6	0.178	0.029	137	1,399.7	144	30.1	4.27
1993-94	133	28.9	0.375	0.125	133	713.5	144	47.6	3.91
1994-95	217	14.8	0.963	0.622	217	14.9	143	60.6	4.86
1995-96	72	79.1	0.986	-	144	29.4	143	19.5	5.00

Krill and salp abundance are averaged field-season values (generally 2 or more cruises). *N* is the number of samples. Chlorophyll-*a* data are the integrated 0–100 m values. Dashes indicate that data are not available. Spawning index is proportion of the mature females in advanced reproductive stages during January–February. Recruitment index is the proportion of 1-year-old krill resulting from this spawning season to the total krill caught next field season. Sea-ice index reflects extent and duration of sea-ice across the Antarctic Peninsula region before austral spring–summer surveys.

Table 2 Salp and krill biomass

Biomass (mg Cm ⁻²)	1994		1995		1996	
	Salps	Krill	Salps	Krill	Salps	Krill
Mean	570.6	314.1	7.8	242.3	20.2	337.3
s.d.	563.2	856.4	16.1	201.1	30.9	756.1
Median	400.5	25.6	1.3	43.5	10.0	72.2
Maximum	3,276.8	4,971.1	75.3	1,545.2	134.2	4,721.0
<i>N</i>	63	63	57	71	72	72
Salp: krill ratio	15.6		0.03		0.1	

Biomass in the Elephant Island area during January 1994, 1995 and 1996. *N* is number of samples. Salp: krill biomass ratios based on median values.

the spring bloom period could deprive krill of sufficient food to support their energy requirements, and therefore retard krill gonadal development and spawning, leading to poor year-class success.

Indices of the spatial and temporal extent of annual sea-ice cover in the Antarctic Peninsula region are tightly correlated with air temperature, allowing the use of the longer time series instead of sea-ice cover¹² (Fig. 2). A warming trend has been documented for the Antarctic Peninsula region since the 1940s (refs 20–22), and a decreased frequency of extensive winter sea-ice conditions has been associated with this trend. The observation that the frequency of cold winters with extensive sea-ice cover has decreased from four out of five years during the middle of this century to only one or two out of five years since the 1970s (refs 12, 13) is confirmed using a regional rather than a local sea-ice index (Fig. 2). Decreased frequency of strong krill year-class success and recruitment, associated with a decreased frequency of extensive winter sea-ice, may be responsible for the significantly lower population sizes of krill observed in the Antarctic Peninsula region since the mid-1980s. Reduced krill population sizes could have a profound effect on their dependent predator populations.

An example of a predator effect may be the decline in the abundance of Adelie penguins (*Pygoscelis adeliae*) among colonies in Admiralty Bay, King George Island. These penguins forage in Bransfield Strait, an area typically populated in the spring and summer months by newly recruited one-year-old krill (ref. 23). Studies since the mid-1970s indicate a significant decline in fledgling survival and a 30% reduction in population size after 1987 (ref. 24).

The large interannual fluctuations in salp and krill populations,

and related shifts in utilization of primary production, may have major implications for the food web. During years favouring strong krill recruitment success and elevated population size, a relatively large proportion of primary production is likely to be incorporated into the krill-based food web and the newly fixed carbon transferred to vertebrate predators or recycled within the upper water column with only occasional transport to deeper water^{6,16}. During periods with massive salp blooms, poor krill recruitment, and declining krill population size, substantially less of the newly fixed carbon may be transferred through krill, with a larger proportion being incorporated into salp fecal pellets and transported to deeper waters where it may be sequestered¹⁶.

Given the current trend for periods of two to three years of above- or below-average winter sea-ice extent²⁵, we can expect continued large variations in the relative levels of dominance by krill and salps within the Antarctic Peninsula region. Associated with increased frequency of years with dense spring–summer salp concentrations is the vastly increased importance of salps as a competitor with krill and as a link in the vertical carbon flux. Occasional years with strong krill recruitment will probably maintain diminished krill population size relative to the 1970s levels. Diminished krill supplies may have long-term negative effects on the upper trophic levels, including commercial harvest by man. These observations, in part, have prompted the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR) to request updated krill abundance surveys to validate or replace the 17-year-old abundance data currently used in the krill management model²⁶. CCAMLR manages the international harvest of Antarctic krill. The current precautionary catch limit for the southwest Atlantic sector is 1.5×10^6 tons yr⁻¹, and for the Indian Ocean sector is 77.5×10^4 tons yr⁻¹. □

Methods

Krill and salp data are derived primarily from December–March net-sampling operations by German expeditions and US Antarctic Marine Living Resources (AMLR) Program surveys. The data in Table 1 generally represent averaged results from two or more cruises each field season. Details on net sampling and sample processing, and considerations of gear selectivity and sampling bias, are presented in refs 11, 14. Krill spawning index is the proportion of females in advanced reproductive stages in January–February¹¹. Krill recruitment index is the proportion of 1-year-old individuals resulting from each spawning season of the total population sampled the next field season (*R*₁); (*R*₂) (the

proportion of 2-year-olds) is used when R_1 is unavailable or has large variance. Recruitment indices are calculated using the maximum-likelihood method^{14,27}. Chlorophyll-*a* data are integrated 0–100 m values. Salp biomass is derived from the internal body length: C relationship¹⁸ (carbon = $0.001336 \times \text{length}^{2.331}$). Krill biomass is derived from the relationships between standard length and wet weight (ref. 28) (weight = $0.00158 \times \text{length}^{3.40}$) and wet weight and carbon (ref. 9) (carbon = $72.77 \times \text{weight}^{1.0242}$). Sea-ice coverage is based on daily ice charts derived from passive microwave imagery by the U.S. National Snow and Ice Data Center²⁹. Annual sea-ice indices for an area of 1,250,000 km² on the west side of the Antarctic Peninsula were determined by numerical integration of monthly estimates of ice cover for each calendar year, and expressed in units of 10⁶ km² months³⁰. The relationship between air temperature and the regional sea-ice index (Fig. 2) is based on mean monthly air temperatures from Palmer Station or from Faraday, adjusted as ref. 21. The regression, temperature = $-6.514 + 0.0505$ (for years between 1947 and 1996) ($F = 13.98$, $P = 0.001$) is similar to that from 1944 to 1987 (ref. 12). The correlation between the regional sea-ice index and air temperature is highly significant ($n = 17$, Kendall's $T = -0.66$, $P < 0.001$).

Received 6 January; accepted 21 April 1997.

- Marr, J. W. S. The natural history and geography of the Antarctic krill (*Euphausia superba* Dana). *Discovery Rep.* **32**, 33–464 (1962).
- Mackintosh, N. A. Distribution of post-larval krill in the Antarctic. *Discovery Rep.* **36**, 95–156 (1973).
- Quetin, L. B. & Ross, R. M. Behavioral and physiological characteristics of the Antarctic krill, *Euphausia superba*. *Am. Zool.* **31**, 49–63 (1991).
- Nicol, S. in *Antarctic Science: Global Concerns* (ed. Hempel, G.) 144–166 (Springer, Berlin, 1994).
- Kock, K.-H. & Shimadzu, Y. in *Southern Ocean Ecology: The BIOMASS Perspective* (ed. El-Sayed, S. Z.) 287–312 (Cambridge Univ. Press, 1994).
- Marchant, H. J. & Murphy, E. J. in *Southern Ocean Ecology: The BIOMASS Perspective* (ed. El-Sayed, S. Z.) 267–285 (Cambridge Univ. Press, 1994).
- Foxton, P. The distribution and life-history of *Salpa thompsoni* Foxton with observations on a related species, *Salpa gerlachei* Foxton. *Discovery Rep.* **34**, 1–116 (1966).
- Casareto, B. E. & Nemoto, T. Salps of the Southern Ocean (Australian sector) during the 1983–84 summer, with special reference to the species *Salpa thompsoni*, Foxton 1961. *Mem. Natl. Inst. Polar Res. (Spec. issue)* **40**, 221–239 (1986).
- Nishikawa, J. et al. Distribution of salps near the South Shetland Islands during austral summer, 1990–1991 with special reference to krill distribution. *Polar Biol.* **15**, 31–39 (1995).
- Gon, O. & Heemstra, C. *Fishes of the Southern Ocean* (J.L.B. Smith Institute of Ichthyology, Grahamstown, 1990).
- Siegel, V. & Loeb, V. Recruitment of Antarctic krill *Euphausia superba* and possible causes for its variability. *Mar. Ecol. Prog. Ser.* **123**, 45–56 (1995).
- Fraser, W. R., Trivelpiece, W. Z., Ainley, D. G. & Trivelpiece, S. G. Increases in Antarctic penguin populations: reduced competition with whales or a loss of sea ice due to global warming? *Polar Biol.* **11**, 525–531 (1992).
- Fraser, W. R. & Trivelpiece, W. Z. in *Foundations for Ecological Research West of the Antarctic Peninsula* (eds Ross, R. M., Hofmann, E. E. & Quetin, L. B.). *American Geophysical Union Research Series* **70**, 257–272 (1996).
- Siegel, V., de la Mare, W. & Loeb, V. Long-term monitoring of krill recruitment and abundance indices in the Elephant Island Area (Antarctic Peninsula). *CCAMLR Science* (in the press).
- Siegel, V. Age and growth of Antarctic euphausiacea (Crustacea) under natural conditions. *Mar. Biol.* **96**, 483–495 (1987).
- Fortier, L., L  Fevre, J. & Legendre, L. Export of biogenic carbon to fish and to the deep ocean: the role of large planktonic micropredators. *J. Plankton Res.* **16**, 809–839 (1994).
- Smetacek, V., Scharek, R. & N  thig, E.-M. in *Antarctic Ecosystems: Ecological Change and Conservation* (eds Kerry, K. R. & Hempel, G.) 103–114 (Springer, Berlin, 1990).
- Huntley, M. E., Sykes, P. E. & Marin, V. Biometry and trophodynamics of *Salpa thompsoni* Foxton (Tunicata: Thaliacea) near the Antarctic Peninsula in austral summer, 1983–1984. *Polar Biol.* **10**, 59–70 (1989).
- Weber, L. H. & El-Sayed, S. Z. Contributions of the net, nano- and picoplankton to the phytoplankton standing crop and primary productivity in the Southern Ocean. *J. Plankton Res.* **9**, 973–994 (1987).
- King, J. C. Recent climate variability in the vicinity of the antarctic peninsula. *Int. J. Climatol.* **14**, 357–369 (1994).
- Smith, R. C., Stammerjohn, S. E. & Baker, K. S. in *Foundations for Ecological Research West of the Antarctic Peninsula* (eds Ross, R. M., Hofmann, E. E. & Quetin, L. B.) *AGU Res. Ser.* **70**, 105–122 (1996).
- Stark, P. Climatic warming in the central Antarctic Peninsula area. *Weather* **49**, 215–220 (1994).
- Siegel, V. in *Antarctic Ocean and Resources Variability* (ed. Sahrhage, D.) 219–230 (Springer, Berlin, 1988).
- Trivelpiece, W. Z. & Trivelpiece, S. G. The impact of global warming on Antarctica's krill-dependent predator populations. *Ecology* (submitted).
- Stammerjohn, S. E. & Smith, R. C. in *Foundations for Ecological Research West of the Antarctic Peninsula* (eds Ross, R. M., Hofmann, E. E. & Quetin, L. B.) *AGU Res. Ser.* **70**, 81–104 (1996).
- CCAMLR Report of the Fifteenth Meeting of the Scientific Committee. SC-CCAMLR-XV (Commiss. Conserv. Antarc. Mar. Living Resources, Hobart, Tasmania, 1996).
- de la Mare, W. K. Estimating krill recruitment and its variability. *CCAMLR Sci.* **1**, 55–69 (1994).
- Kils, U. Swimming behaviour, swimming performance, and energy balance of Antarctic krill *Euphausia superba*. *BIOMASS Sci. Ser.* **3**, 1–233 (1981).
- National Snow and Ice Data Center, University of Colorado at Boulder (nsidc@kryos.colorado.edu).
- Hewitt, R. P. Areal and seasonal extent of sea ice cover off the northwestern side of the Antarctic Peninsula: 1979 through 1996. *CCAMLR Sci.* (in the press).

Acknowledgements. We thank S. Bros, G. Cailliet, D. Outram, R. Phleger, E. Shulenberger, M. White and K. Wishner for comments, and S. Nicol for helping to improve the manuscript. This work was supported by the National Oceanic and Atmospheric Administration U.S. Antarctic Marine Living Resources Program. Additional support came from the National Science Foundation (W.F., W.T. and S.T.) and the Institut fur Seefischerei (V.S.).

Correspondence and requests for materials should be addressed to V.L. (e-mail: loeb@mlml.calstate.edu).

Corticofugal modulation of frequency processing in bat auditory system

Yunfeng Zhang, Nobuo Suga & Jun Yan

Department of Biology, Washington University, 1 Brookings Drive, St Louis, Missouri 63130, USA

Auditory signals are transmitted from the inner ear through the brainstem to the higher auditory regions of the brain. Neurons throughout the auditory system are tuned to stimulus frequency, and in many auditory regions are arranged in topographical maps with respect to their preferred frequency. These properties are assumed to arise from the interactions of convergent and divergent projections ascending from lower to higher auditory areas¹; such a view, however, ignores the possible role of descending projections from cortical to subcortical regions^{2–10}. In the bat auditory system, such corticofugal connections modulate neuronal activity to improve the processing of echo-delay information^{11,12}, a specialized feature. Here we show that corticofugal projections are also involved in the most common type of auditory processing, frequency tuning. When cortical neurons tuned to a specific frequency are inactivated, the auditory responses of subcortical neurons tuned to the same frequency are reduced. Moreover, the responses of other subcortical neurons tuned to different frequencies are increased, and their preferred frequencies are shifted towards that of the inactivated cortical neurons. Thus the corticofugal system mediates a positive feedback which, in combination with widespread lateral inhibition, sharpens and adjusts the tuning of neurons at earlier stages in the auditory processing pathway.

The moustached bat, *Pteronotus parnellii*, emits orientation sounds (biosonar pulses) that contain a long constant-frequency component (~61 kHz) suited for velocity (Doppler) measurement. The bat uses velocity information carried by the 61-kHz component for hunting flying insects^{13,14}. Accordingly, its auditory system, from the periphery to the cortex, contains many neurons that are sharply tuned to this frequency for fine frequency analysis^{15,16}. The sharp frequency-tuning curves in the central auditory system are sculpted by lateral inhibition^{17–21}. The area of the primary auditory cortex that is tuned to ~61 kHz receives input from part of the thalamus²² and has a frequency axis^{16,17} like that of the primary auditory cortex of rats, cats and monkeys²³.

We studied the effect of inactivation of the corticofugal system on the auditory responses of subcortical neurons (32 thalamic and 27 collicular neurons) that are tuned to frequencies between 60.47 and 62.30 kHz. Lidocaine (a local anaesthetic) always had an effect on the subcortical neurons when applied to cortical neurons tuned to these frequencies but had no effect when applied to cortical neurons tuned to frequencies outside this range (such as 34, 57 or 84 kHz). The effect of Lidocaine when the 'best frequency' of a subcortical neuron and that of inactivated cortical neurons were 'matched' (were within ± 0.20 kHz) was different from when they were 'unmatched' (different by more than 0.20 kHz).

Cortical inactivation decreased the auditory responses of matched subcortical neurons (Fig. 1A,b) without shifting their frequency–response curves (Fig. 1A,d). In contrast, cortical inactivation increased the auditory responses of unmatched subcortical neurons and shifted the best frequency towards that of the inactivated cortical neurons. Recovery of normal tuning took 1.5–3.2 h; such an increase and recovery is shown for two unmatched thalamic neurons that had best frequencies 0.30 kHz lower (Fig. 1B) or 0.68 kHz higher (Fig. 1C) than the inactivated cortical neurons. The frequency–response curve shown in Fig. 1B,d (filled circles)

indicates that the increase in the response of the thalamic neuron occurred mostly for frequencies higher than the best frequency (60.67 kHz). Because of this frequency-dependent increase, the frequency-response curve was shifted towards higher frequencies, that is, towards the best frequency of the inactivated cortical neurons ($P < 0.002$). When the thalamic best frequency was higher than the cortical best frequency, the shift in the thalamic best frequency was toward lower frequencies (Fig. 1C,d, filled circles; $P < 0.004$).

Thalamic and collicular neurons normally showed large variations in the shape of frequency-tuning curves (Fig. 2A–D). Cortical inactivation shifted the frequency-tuning curves of unmatched subcortical neurons, together with their best frequencies, along the frequency axis, but evoked little change in either minimum threshold or best amplitude. The shifted curves returned to the control condition in 1.5–3.2 h. This parallel shift and recovery in the frequency-tuning curves are shown for two thalamic (Fig. 2A,B) and two collicular (Fig. 2C,D) neurons that had best frequency lower (Fig. 2A,D) or higher (Fig. 2B,C) than those of inactivated cortical neurons. Although the shift was parallel, a small amount of broadening in tuning-curve width was noticed in some subcortical neurons (for example, at 40–60 dB sound pressure level in Fig. 2C). The amount of broadening varied with stimulus amplitude: the

higher the stimulus level, the greater the broadening (Fig. 2E).

After cortical inactivation, thalamic neurons showed a larger change in response magnitude than did collicular neurons. The decrease in the response of matched neurons was 2.6 times larger for the thalamic neurons ($54.2 \pm 14.5\%$ (mean $\pm$ s.d.), $n = 4$) than for the collicular neurons ($20.8 \pm 8.90\%$, $n = 4$) (Fig. 3A,C, open triangles; $P < 0.05$). The increase in the response of unmatched neurons measured at the best frequencies shifted by the cortical inactivation was 1.8 times larger for the thalamic neurons ($104 \pm 59.5\%$, $n = 28$) than for the collicular neurons ($59.0 \pm 39.5\%$, $n = 23$) (Fig. 3A,C, filled circles; $P < 0.01$).

After cortical inactivation, best frequencies did not shift for subcortical matched neurons (Fig. 3B,D, open triangles), but shifted for subcortical unmatched neurons: the larger the difference in best frequency, the larger the shift (Fig. 3B,D, filled circles). The rate of shift per difference in best frequency was 0.33 for the thalamic neurons and 0.18 for the collicular neurons. Therefore, the magnitude of best-frequency shift was 1.8 times larger for the thalamic neurons than for the collicular neurons ($P < 0.05$). These data indicate that the decrease or increase in response and the shift in frequency tuning of the subcortical neurons occur in both the colliculus and the thalamus through the corticocollicular and corticothalamic projections.

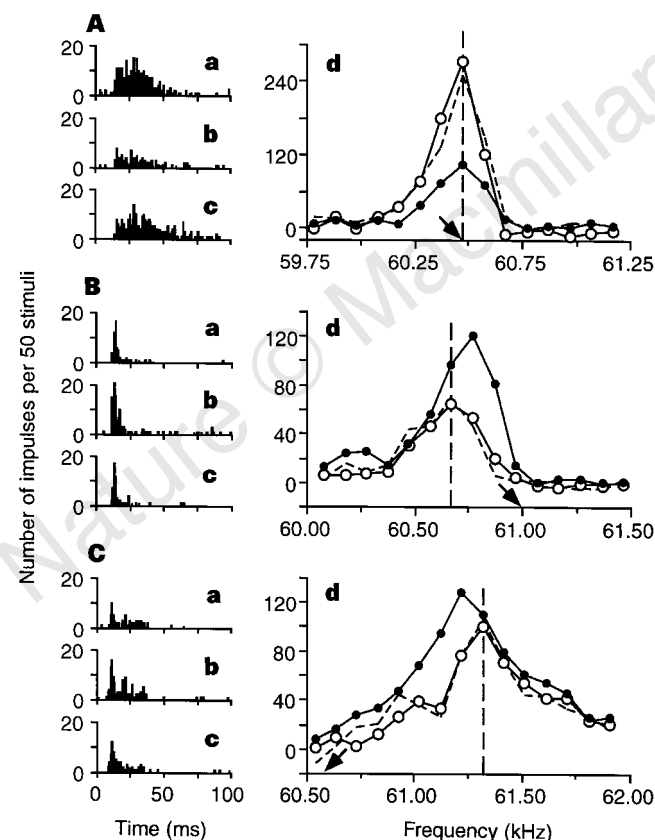


Figure 1 Auditory responses (a–c) and frequency-response curves (d) of three single thalamic neurons obtained before (a; control, open circles in d), during (b; Lidocaine application, 90 nl, 1.0%, filled circles in d) and after (c; recovery, broken line in d) a focal inactivation of cortical neurons. The best frequencies of cortical and thalamic neurons paired for the experiments, respectively, were: **A**, 60.47 and 60.47 kHz; **B**, 60.97 and 60.67 kHz, and **C**, 60.64 and 61.32 kHz. The effects of the inactivation of the cortical neurons on the thalamic neurons were different depending on the relationship in best frequency between the cortical and thalamic neurons. The vertical broken lines and arrows indicate the best frequencies of the thalamic and cortical neurons, respectively.

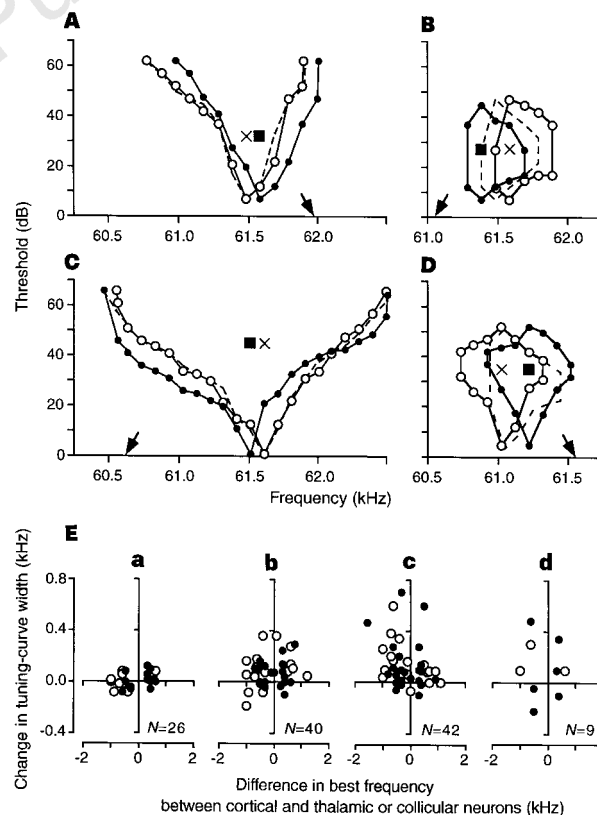


Figure 2 Changes in the frequency-tuning curves of thalamic (**A** and **B**) and collicular neurons (**C** and **D**) by a focal inactivation of cortical neurons. The best frequencies of the inactivated cortical neurons are indicated by the arrows. The curves were measured before (control, open circles), during (filled circles) and after (recovery, broken lines) the cortical inactivation. The tuning curves shift towards the best frequencies of inactivated cortical neurons. Crosses and squares indicate the best amplitudes measured before and during the cortical inactivation, respectively. **E**, Changes in the widths of frequency-tuning curves at 10 (a), 30 (b), 50 (c) and 70 dB (d) above minimum threshold evoked by a focal inactivation of cortical neurons. Results are for thalamic (filled circles) and collicular (open circles) neurons.

In our experiments, the number of matched subcortical neurons studied was small ($n = 8$). However, the corticofugal effects are strikingly consistent and are, without exception, completely opposite to those on the 51 unmatched subcortical neurons (Fig. 3A,C). Furthermore, the data from the matched neurons (no best-frequency shifts) and those from the unmatched neurons (best-frequency shifts) fit the same linear function as shown by the regression lines in Fig. 3, which cross the zero-shift line at a zero best-frequency difference. Therefore, the results justify our conclusion that, in the normal condition, cortical neurons augment the responses of matched subcortical neurons, but suppress those of unmatched subcortical neurons and shift their frequency-tuning curves away from those of the cortical neurons.

The frequency axis in the cortical area tuned to 61 kHz is radial. Best frequency changes at a rate of $\sim 3.3 \text{ kHz mm}^{-1}$ for frequencies between 61.0 and 61.5 kHz along the rostrocaudal axis of this cortical area^{16,17,24}. This corresponds to a rate of $\sim 66 \text{ Hz per minicolumn}$ (a minicolumn is a slab formed by neurons tuned to an identical best frequency), because each minicolumn is $\sim 20 \mu\text{m}$ wide. The best frequencies of cortical neurons that had a positive feedback effect on a subcortical neurons were within a range of $\pm 0.2 \text{ kHz}$ of the best frequency of the subcortical neuron. This indicates that a subcortical neuron receives positive feedback from cortical neurons in a maximum of six minicolumns. Since Lidocaine applied to the cortex would spread to some extent, it is most likely that a subcortical neuron receives positive feedback from only one or a few minicolumns.

Our data indicate that individual subcortical and, perhaps, cortical neurons have multiple excitatory inputs that differ slightly in best frequency but are otherwise similar in frequency and amplitude tuning. Second, our data indicate that cortical neurons select particular excitatory inputs through a highly focused positive feedback that is associated with widespread lateral inhibition; that

is, they perform 'egocentric selection' through the corticofugal system. Third, egocentric selection shifts the frequency-tuning curves of unmatched subcortical neurons, and sharpens some of these curves. Fourth, egocentric selection enhances the contrast in the neural representation of auditory information. Finally, our data indicate that egocentric selection is one of the fundamental functions of the auditory cortex, because it is found for frequency-domain processing, which is shared by all higher vertebrates, as well as for time (echo-delay)-domain processing which may be unique to the auditory system of the bat.

The effects of a 100-nA, 0.2-ms electrical stimulation of the cortex on subcortical neurons lasted $\sim 2 \text{ h}$ when it was delivered to the cortex at a rate of 6 stimulations per second for 11 min (ref. 12). The effects of Lidocaine were apparent for 1.5–3.2 h even though only a small amount was applied. Egocentric selection works with a slow time course. When an identical signal is received repeatedly by an animal, egocentric selection will increase the neural representation of this frequently occurring signal in the colliculus, thalamus and cortex. When a signal is received only rarely by the animal, however, the neural representation will decrease for this signal. Therefore, we hypothesize that egocentric selection is involved in long-term changes in the overall functional organization of the central auditory system^{25–27}.

Almost all auditory neurons are tuned to particular frequencies. Therefore, all signal processing, including ranging and sound localization, is presumably modulated in the frequency domain by the corticofugal projections originating in the primary auditory cortex. In addition, signal processing in other domains, such as echo delay, are modulated more specifically by the corticofugal projections originating from specialized cortical area. The descending system originating from the cortex eventually ends in the cochlear hair cells. It is not yet known whether egocentric selection occurs in the subcollicular nuclei and cochlea. □

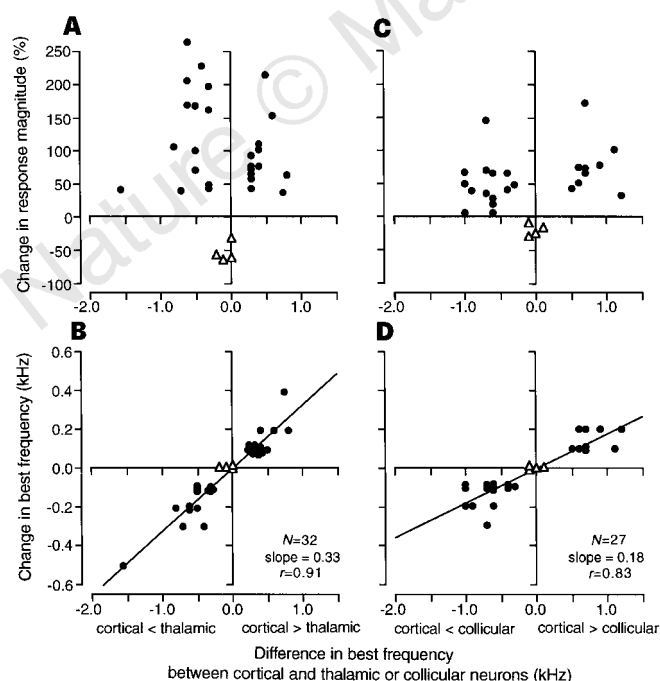


Figure 3 Changes in response magnitude (A and C), and best frequency (B and D) of thalamic (A and B) and collicular neurons (C and D) evoked by a focal inactivation of cortical neurons. Triangles and filled circles represent the data obtained from matched and unmatched subcortical neurons, respectively. The filled circles in A and C represent the increase in response magnitude at the best frequencies shifted by cortical inactivation. The regression line, its slope, and correlation coefficient (r) are shown in B and D.

Methods

General. The preparation of animals and recording of action potentials were as described²⁸. In four moustached bats not anaesthetized, the best frequencies of multiple neurons were first measured at 4–5 loci in the area of the primary auditory cortex tuned to 60.47–62.30 kHz (ref. 24). The best frequency of a single neuron in the ventral division of the medial geniculate body in the thalamus or the dorsoposterior division of the inferior colliculus in the midbrain was measured. To inactivate one of the cortical loci where best frequencies were measured, 90 nl of 1.0% Lidocaine (local anaesthetic) was injected into the $\sim 900\text{-}\mu\text{m}$ -thick cortex at a depth of 600–700 μm by using a mechanical micro-injection unit. Responses of the single subcortical neuron to tone bursts of different frequencies and amplitudes were collected with a computer before, during and after the cortical inactivation. A t -test was used to establish whether the difference in auditory response was statistically significant before and after Lidocaine or between thalamic and collicular neurons.

Acoustic stimuli. To measure a frequency–response curve, a computer-controlled frequency scan was presented 50 times at a best amplitude excite to a given neuron. The frequency scan consisted of 21 time blocks: in the first 20 blocks, the frequency of a tone burst lasting 23 ms was changed by 0.1-kHz steps across the best frequency of a given neuron; in the 21st (last) block, no stimulus was presented. The duration of each block was 200 ms. To measure a frequency-tuning curve, the frequency scan consisted of 33 time blocks, in the first 32 of which frequency was changed by 0.1-kHz steps. This frequency scan was presented 5 times at a given amplitude, and after every 5 frequency scans the amplitude of a tone burst was changed from 101 dB to 1 dB in 5-dB steps. Thus the frequency–amplitude scan consisted of 33×21 blocks.

Data acquisition. The responses of a neuron during the frequency or frequency–amplitude scan were continuously monitored as rasters and as an array of peristimulus–time histograms displayed on a computer screen. Action potentials were also continuously monitored by comparing incoming action potentials with the template of an action potential recorded at the beginning of the continuous recording with a digital storage oscilloscope. Data were used for off-line analysis as long as the shape of action potentials matched the template.

Data processing. The criterion for threshold was the boundary between a block showing one response (action potential) per five trials and a block showing no response per five trials. In unanaesthetized bats the response magnitude and background discharge rate of thalamic and collicular neurons fluctuated to some extent. Therefore, the threshold was determined with additional criteria. If a given block showed at least one response per five trials, it was interpreted to be within an excitatory area demarcated by a tuning curve. If it showed no response per five trials, but the adjacent blocks, away from the excitatory area, showed at least one response per five trials, the given block was put within the excitatory area. Because auditory responses always occurred within a time window of 50 ms with relatively constant latencies ranging between 5 and 12 ms, and background discharge rate was low ($0.58 \pm 0.32 \text{ s}^{-1}$ for 17 collicular neurons and $0.85 \pm 0.43 \text{ s}^{-1}$ for 16 thalamic neurons), auditory responses were easily distinguished from background discharges.

We established the criteria for a shift in frequency tuning as follows. If frequency-response or frequency-tuning curves shifted by cortical inactivation with Lidocaine did not recover to more than 50% of their earlier value they were excluded from our analysis. The fact that almost all curves shifted by Lidocaine recovered by more than 50% indicates that the shift was significant. To ensure that the shift was significant we used the following procedure. For individual frequency-response curves, a weighted average frequency (best frequency) was calculated for the summed responses to 5 consecutive frequency scans. We obtained 10 such values to calculate the mean $\pm$ s.d. We used a two-tailed, paired *t*-test to determine whether or not the weighted average frequencies (best frequencies) obtained for control and Lidocaine conditions were significantly different (at $P < 0.01$). Frequency-tuning curves were based upon the threshold measurement, so s.d. could not be calculated for each data point. We therefore applied the following criterion: shifts in best frequencies and frequency-tuning curves evoked by Lidocaine were considered to be significant if the shift in best frequency was accompanied by shifts of more than 75% of the data points in the same direction as that of the best-frequency shift.

Received 31 December 1996; accepted 25 April 1997.

1. Suga, N. Cortical computational maps for auditory imaging. *Neural Networks* **3**, 3–21 (1990).
2. Huffman, R. F. & Henson, O. W. Jr The descending auditory pathway and acousticomotor system: connections with the inferior colliculus. *Brain Res.* **15**, 295–323 (1990).
3. Andersen, R. A. *et al.* The topographic organization of corticocollicular projections from physiologically identified loci in the AI, AII, and anterior auditory cortical field of the cat. *J. Comp. Neurol.* **191**, 479–494 (1980).
4. Herbert, H. *et al.* Topography of projections from the auditory cortex to the inferior colliculus in the rat. *J. Comp. Neurol.* **304**, 103–122 (1991).
5. Massopust, L. C. Jr & Ord, J. M. Auditory organization of the inferior colliculus in the cat. *Exp. Neurol.* **6**, 465–477 (1962).
6. Andersen, P. *et al.* Corticofugal facilitation of thalamic transmission. *Brain Behav. Evol.* **6**, 170–184 (1972).
7. Ryugo, D. K. & Weinberger, N. M. Corticofugal modulation of the medial geniculate body. *Exp. Neurol.* **51**, 377–391 (1976).
8. Mitani, A. *et al.* Effects of stimulation of the primary auditory cortex upon colliculogeninate neurons in the inferior colliculus of the cat. *Neurosci. Lett.* **42**, 185–189 (1983).
9. Sun, X. *et al.* Corticofugal influences on the responses of bat inferior colliculus to sound stimulation. *Brain Res.* **495**, 1–8 (1989).
10. Villa, A. P. E. *et al.* Corticofugal modulation of the information processing in the auditory thalamus of the cat. *Exp. Brain Res.* **86**, 506–517 (1991).
11. Suga, N. *et al.* in *Active Hearing* (eds Ottonson, A. F. D. & Ulfendahl, M.) 13–30 (Pergamon, London, 1995).
12. Yan, J. & Suga, N. Corticofugal modulation of time-domain processing of biosonar information in bats. *Science* **273**, 1100–1103 (1996).
13. Johnson, R. A. *et al.* Detection of insect wing beats by the bat, *Pteronotus parnellii*. *J. Acoust. Soc. Am.* **55**, S53 (1974).
14. Goldman, L. J. & Henson, O. W. Jr Prey recognition and selection by the constant frequency bat, *Pteronotus p. parnellii*. *Behav. Ecol. Sociobiol.* **2**, 411–419 (1977).
15. Suga, N. *et al.* Peripheral specialization for fine analysis of Doppler-shifted echoes in the auditory system of the “CF-FM” bat *Pteronotus parnellii*. *Exp. Biol.* **63**, 161–192 (1975).
16. Suga, N. & Jen, P. H. S. Disproportionate tonotopic representation for processing species-specific CF-FM sonar signals in the mustached bat auditory cortex. *Science* **194**, 542–544 (1976).
17. Suga, N. & Manabe, T. Neural basis of amplitude-spectrum representation in auditory cortex of the mustached bat. *J. Neurophysiol.* **47**, 225–255 (1982).
18. Suga, N. & Tsuzuki, K. Inhibition and level-tolerant frequency-tuning curves in the auditory cortex of the mustached bat. *J. Neurophysiol.* **53**, 1109–1145 (1985).
19. Yang, L. *et al.* GABAergic circuits sharpen tuning curves and modify response properties in the mustached bat inferior colliculus. *J. Neurophysiol.* **68**, 1760–1774 (1992).
20. Suga, N. *et al.* Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. *J. Neurophysiol.* **77**, 2098–2114 (1997).
21. Suga, N. Sharpening of frequency tuning by inhibition in the central auditory system: Tribute to Yasuji Katsuki. *Neurosci. Res.* **21**, 287–299 (1995).
22. Olsen, J. F. thesis, Washington University (1986).
23. Winer, J. A. in *The Mammalian Auditory Pathway: Neuroanatomy* (eds Webster, D. B., Popper, A. N. & Fay, R. R.) 222–410 (Springer, New York, 1992).
24. Suga, N. *et al.* The personalized auditory cortex of the mustached bat: adaptation for echolocation. *J. Neurophysiol.* **58**, 643–654 (1987).

25. Edeline, J. M. & Weinberger, N. M. Subcortical adaptive filtering in the auditory system: associative receptive field plasticity in the dorsal medial geniculate body. *Behav. Neurosci.* **105**, 154–175 (1991).
26. Edeline, J. M. & Weinberger, N. M. Associative returning in the thalamic source of input to the amygdala and auditory cortex: receptive field plasticity in the medial division of the medial geniculate body. *Behav. Neurosci.* **106**, 81–105 (1992).
27. Edeline, J. M. *et al.* Rapid development of learning-induced receptive field plasticity in the auditory cortex. *Behav. Neurosci.* **107**, 539–551 (1993).
28. Suga, N. *et al.* Specificity of combination-sensitive neurons for processing of complex biosonar signals in the auditory cortex of the mustached bat. *J. Neurophysiol.* **49**, 1573–1626 (1983).

Acknowledgements. We thank the Natural Resource Conservation Authority and the Ministry of agriculture of Jamaica for allowing us to collect and export the mustached bats used in our research. We thank S. Kuwada, W. E. O'Neill, J. F. Olsen, D. C. Fitzpatrick and A. Kadir for helpful comments.

Correspondence and requests for materials should be addressed to N.S. (e-mail: suga@biodec.wustl.edu).

Congenital leptin deficiency is associated with severe early-onset obesity in humans

Carl T. Montague^{*†}, I. Sadaf Farooqi^{*†‡}, Jonathan P. Whitehead^{*†}, Maria A. Soos^{*†}, Harald Rau^{*†}, Nicholas J. Wareham[§], Ciaran P. Sewter^{*†}, Janet E. Digby^{*†}, Shehla N. Mohammed^{||}, Jane A. Hurst[†], Christopher H. Cheetham[#], Alison R. Earley[#], Anthony H. Barnett[☆], Johannes B. Prins^{*†} & Stephen O'Rahilly^{*†}

University of Cambridge, Departments of ^{*} Medicine, [‡] Clinical Biochemistry and [§] Community Medicine, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QR, UK

^{||} South Thames Regional Genetics Centre (East), Guy's Hospital, London SE1 9RT, UK

[¶] Oxford Regional Genetics Service, Churchill Hospital, Oxford OX3 7LJ, UK

[#] Wycombe General Hospital, Queen Alexandra Road, High Wycombe, Buckinghamshire HP11 2TT, UK

[☆] Department of Medicine, University of Birmingham and Birmingham Heartlands Hospital, Birmingham B9 5SS, UK

[†] These authors contributed equally to this study.

The extreme obesity of the *obese (ob/ob)* mouse is attributable to mutations in the gene encoding leptin¹, an adipocyte-specific secreted protein which has profound effects on appetite and energy expenditure. We know of no equivalent evidence regarding leptin's role in the control of fat mass in humans. We have examined two severely obese children who are members of the same highly consanguineous pedigree. Their serum leptin levels were very low despite their markedly elevated fat mass and, in both, a homozygous frame-shift mutation involving the deletion of a single guanine nucleotide in codon 133 of the gene for leptin was found. The severe obesity found in these congenitally leptin-deficient subjects provides the first genetic evidence that leptin is an important regulator of energy balance in humans.

In 1994, two different strains of *ob/ob* mice were reported to have defects in the gene encoding leptin, a previously unknown secreted fat-cell product¹. Leptin is thought to act primarily at the hypothalamus, where it has effects on appetite, energy expenditure and neuroendocrine axes^{2–4}. Treatment of *ob/ob* mice with biosynthetic leptin corrects all of their phenotypic abnormalities^{5–7}. Further, the administration of large amounts of leptin to normal rats and mice markedly reduces body fat stores, suggesting that, at least in rodents, leptin may influence body fat mass across a range of serum concentrations⁸. A role for leptin deficiency in human obesity has been considered but no pathogenic mutations in the gene that encodes leptin have previously been found in obese humans^{9,10}. Despite considerable evidence that genetic factors contribute to human obesity¹¹, no mutations in any gene have been reported to cause obesity in humans. We have studied two related children with

extreme obesity, and our observations indicate that their obesity is due to a congenital deficiency in the production of leptin.

The two children, here referred to as Ob1 and Ob2, are cousins within a highly consanguineous family of Pakistani origin. Although of normal weight at birth, both children suffered from severe, intractable obesity from an early age. The children have no additional clinical features to suggest that they might have a pleiotropic genetic syndrome associated with obesity, such as

Alstrom's or Prader-Willi syndrome. Previous investigations have included normal karyotypes and normal computer tomography (CT) of the brain of Ob1. Each child has two siblings of normal weight, and none of their parents are morbidly obese. Informed parental consent was obtained for all studies, and ethical permission was granted by the Cambridge Local Research Ethics Committee.

To examine whether a mutation in the gene that encodes leptin might underlie severe obesity in subject Ob2, the nucleotide

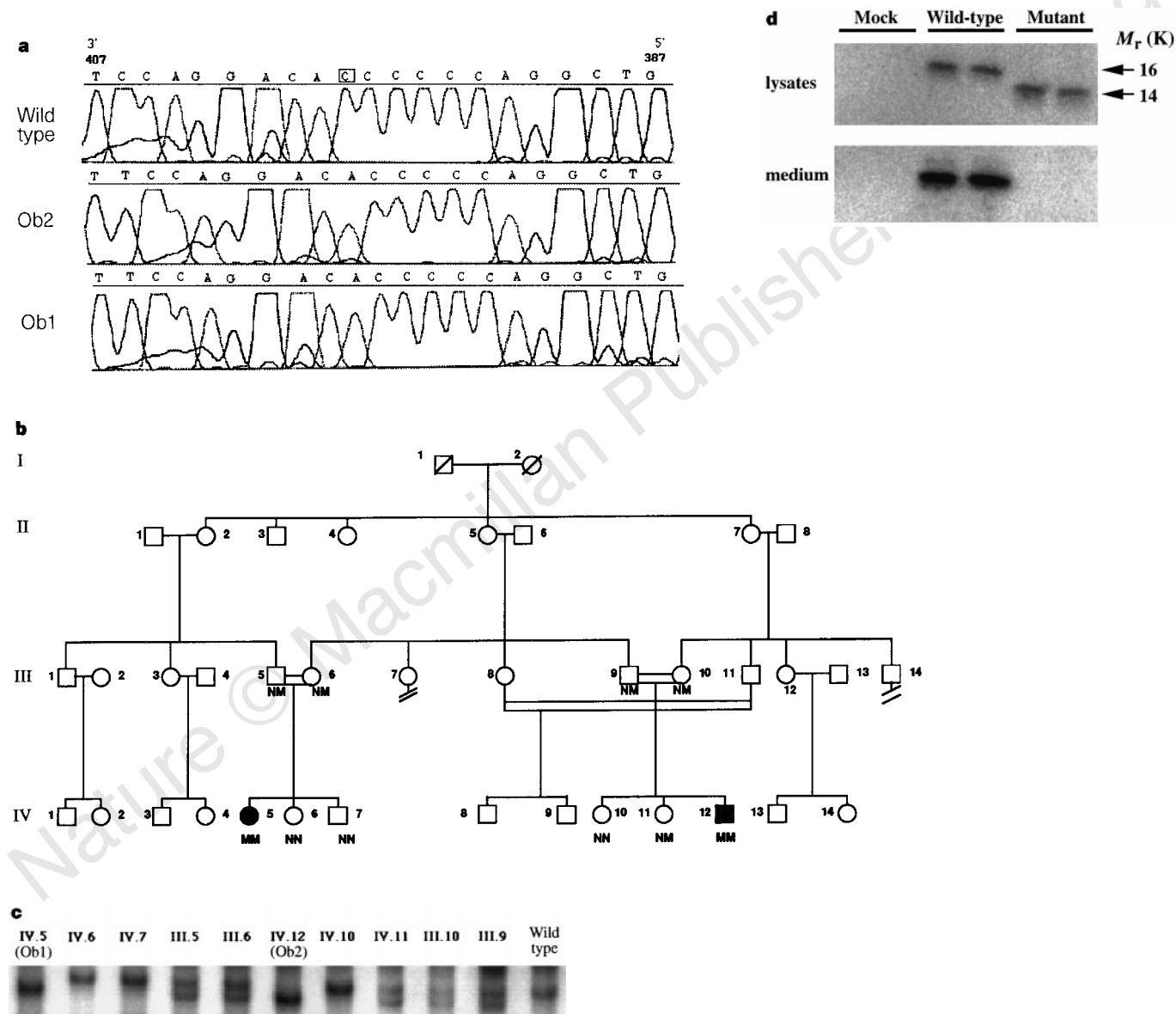


Figure 1 a, Detection of a homozygous frame-shift mutation in the leptin genes of probands Ob1 and Ob2. Adipocyte RNA was isolated from a needle biopsy of subcutaneous adipose tissue obtained from Ob2. RNA was reverse-transcribed and the nucleotide sequence encoding the leptin protein was amplified by PCR and subcloned, and four subclones were sequenced. The nucleotide sequence provided by the automated sequencing apparatus corresponds to the reverse complement of the coding sequence. In all four clones of Ob2's leptin cDNA, five, rather than the expected six, guanine nucleotides were present between nucleotides 393 and 398. The same region of the leptin coding sequence was amplified from genomic DNA of Ob1 and an unrelated control subject. The leptin gene of the control subject has six guanines (G) in the appropriate position (shown as cysteines (C), as the reverse complement strand was sequenced), whereas Ob1 is homozygous for the deleted guanine. **b, c**, Individuals with severe early-onset obesity are indicated by black symbols. The genotype, if known, is indicated below the symbol: N, normal; M, leptin mutation. Clinically unaffected

family members are not homozygotes for the frame-shift mutation. The inheritance of the leptin frame-shift mutation in available first-degree relatives of Ob1 and Ob2 was examined by PCR-SSCP (**c**). The mutant sequence resulted in a readily detectable mobility shift, which was clearly seen in homozygous form in Ob1 and Ob2. As expected, all four parents were heterozygotes. Of the siblings of the probands, three were homozygous for the wild-type sequence, and one was a heterozygote. These results were confirmed by direct sequencing (data not shown). **d**, The secretion of mutant leptin is impaired. CHO cells were transiently transfected with expression vectors encoding either mutant or wild-type leptin. Leptin immunoreactivity was examined in cell lysates (top) and in the medium (bottom) by immunoprecipitation, SDS-PAGE and western blotting. With wild-type leptin, the expected band at 16K is seen in both the cell lysate and the medium, whereas, in cells transfected with mutant leptin, no leptin is seen in the medium, although a truncated 14K species is detectable in the lysate.

sequence of leptin cDNA (Genbank accession no. U18915) was examined. Leptin cDNA was generated from reverse-transcribed RNA isolated from 20 mg of subcutaneous adipose tissue obtained by needle biopsy. Polymerase chain reaction (PCR) products were subcloned and four independent clones sequenced. All clones from subject Ob2 contained a frame-shift mutation in the leptin-coding region consisting of the deletion of a single guanine nucleotide normally present in codon 133 (see Fig. 1a for the reverse complement sequence). Both Ob1 and Ob2 were confirmed to be homozygous for this deletion by direct sequencing of PCR products of the leptin gene amplified from their genomic DNA (Fig. 1a). The mutation disrupts the reading frame of the leptin gene, leading to the introduction of 14 aberrant amino acids after Gly 132 in the native leptin polypeptide, followed by a premature stop codon.

To examine the co-segregation of the leptin gene mutation with body fat content in this pedigree, 176 base pairs of the leptin gene encompassing the site of the deleted guanine nucleotide were examined by PCR-single strand conformation polymorphism (SSCP) analysis in the probands and their available first-degree relatives using genomic DNA isolated from blood leukocytes as a template. The mutation resulted in a readily detectable mobility shift, which was seen in homozygous form in both Ob1 and Ob2

(Fig. 1c). As expected, all four parents were heterozygotes. Of the four siblings of the probands, one was a heterozygote and three were homozygous for the wild-type sequence (Fig. 1b, c). The genotype of all family members was confirmed by nucleotide sequencing. The mutation was not found by SSCP analysis in 108 alleles of UK control subjects of Pakistani origin.

Chinese hamster ovary (CHO) cells were transiently transfected with expression vectors encoding either wild-type or mutant leptin cDNA. Although 16K wild-type leptin was readily detectable in the medium by immunoprecipitation, SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting, no secreted mutant leptin was detected (Fig. 1d). The absence of mutant leptin in the medium was not due to failure of the antibodies to recognize the mutant leptin because, when cells expressing the mutant leptin were lysed, a 14K truncated species was readily detected (Fig. 1d). It is therefore likely that the frame-shift mutation results in a form of leptin which is not targeted normally for secretion.

To examine whether homozygosity for this frame-shift mutation was associated with abnormalities in circulating leptin, serum leptin levels were measured by radio-immunoassay in Ob1 and Ob2 and found to be close to the limits of detection of this assay at 1.0 and 0.7 ng ml⁻¹, respectively. The mean (±s.d.) serum leptin levels found in 16 normal prepubertal children was 8 ± 4.5 ng ml⁻¹, and in 30 adults was 23 ± 10 ng ml⁻¹. Because serum leptin levels show a strong positive correlation with indices of obesity¹² (see Fig. 2), the finding of barely detectable levels of serum leptin in these two extremely obese children was notable. As the leptin levels measured by radio-immunoassay were very low, serum leptin levels in Ob1 and Ob2 were examined further using a highly sensitive solid-phase sandwich enzyme immunoassay. Serum leptin levels in Ob1 and Ob2 were 0.11 and 0.38 ng ml⁻¹, respectively, again close to the detection limit of this assay (0.07 ng ml⁻¹). Finally, no immunoreactive leptin was detected in the serum of Ob2 by western blotting. A band identical to the recombinant standard 16K was readily detected in serum from all four heterozygote adults, but no immunoreactive material of lower molecular weight was seen in the serum of these subjects (M. Nicolson, data not shown). As the antibodies used to detect leptin in all three assay systems were polyclonal and raised to the entire protein, it is possible that we have underestimated the circulating concentrations of the mutant leptin

Table 1 Endocrine biochemistry of subjects Ob1 and Ob2

	Ob1	Ob2	Reference range
fasting glucose	4.7	4.4	3.5–5.5 mmol l ⁻¹
insulin	158	46	0–60 pmol l ⁻¹
proinsulin	21	8.1	0–5 pmol l ⁻¹
32:33 split proinsulin	39	11	<13 pmol l ⁻¹
9:00 ACTH	58	28	5–50 ng l ⁻¹
9:00 cortisol	420	156	280–650 nmol l ⁻¹
TSH	4.80	4.6	0.4–4 mU l ⁻¹
free T4	11.4	12.4	9–20 pmol l ⁻¹
FSH	0.8	0.2	prepubertal 0.6–3.4 U l ⁻¹
LH	<0.2	<0.2	prepubertal 0.6–1.7 U l ⁻¹
oestradiol	<20	—	prepubertal <60 pmol l ⁻¹
testosterone	—	<0.2	prepubertal <0.7 nmol l ⁻¹

Levels of testosterone, oestradiol, luteinizing hormone (LH) and follicle-stimulating hormone (FSH) are all consistent with the prepubertal state. Ob1 has marked fasting hyperinsulinaemia and a mild elevation in 9:00 levels of adrenocorticotrophic hormone (ACTH) and thyrotrophic hormone (TSH).

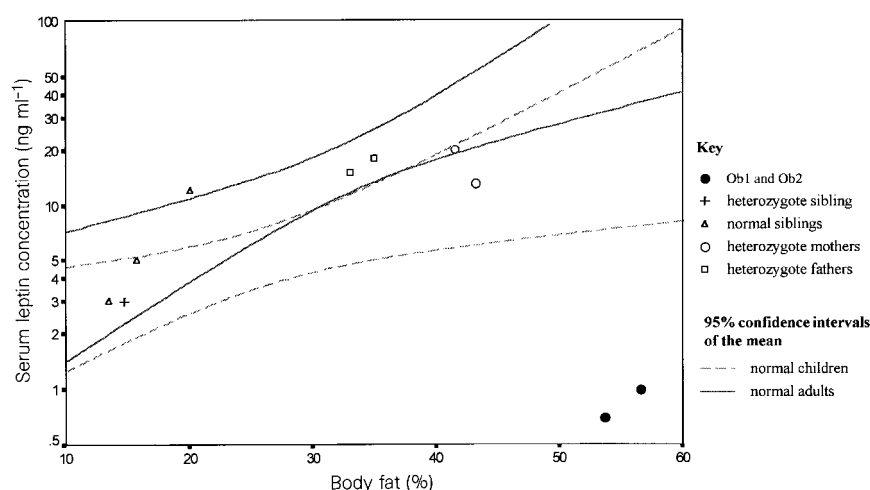


Figure 2 Serum leptin concentrations in Ob1 and Ob2 deviate markedly from the normal relationship between percentage body fat and serum leptin. Serum leptin concentrations were measured by radio-immunoassay in the probands, their relatives, adult and prepubertal control subjects. As expected, serum leptin levels in both adult and prepubertal controls correlated positively with estimates of percentage body fat. The serum leptin levels of Ob1 and Ob2 were markedly reduced, being the lowest recorded in this study at 1.0 and 0.7 ng ml⁻¹,

respectively. Moreover, the finding of low circulating leptin in Ob1 and Ob2 represents a marked deviation from the expected relationship between percentage body fat and serum leptin. This relationship was not markedly different in the relatives who were heterozygous for the frame-shift mutation compared either to family members who were homozygous for the wild-type sequence or to unrelated control subjects.

in Ob1 and Ob2. Even if this were the case, it is unlikely that the mutant leptin would have any biological activity, as it lacks the carboxy-terminal cysteine that is required for intra-chain disulphide bonding. Mutation of this cysteine has been reported to render the protein biologically inactive¹³. The coexistence in Ob1 and Ob2 of low levels of immunoreactive serum leptin, gross obesity and a homozygous frame-shift mutation of the leptin gene are, therefore, likely to be causally linked.

The frame-shift mutation may adversely affect serum leptin levels in several ways. First, premature stop codons frequently result in instability of mRNA and a low level of steady-state mRNA¹⁴. Second, the introduction of an aberrant stretch of amino-acid sequence into the native peptide may disrupt the intracellular targeting and secretion of leptin, as is suggested by the transfection experiments in CHO cells.

The *ob/ob* mouse is of normal weight at birth, but rapidly exhibits hyperphagia and is characterized by several phenotypic abnormalities and severe obesity. We studied the phenotype of subjects Ob1 and Ob2 to determine the degree of similarity between the murine and human syndromes of congenital leptin deficiency. Given the ethical constraints on the study of prepubertal children and our concern to conform to parental wishes, phenotypic characterization was restricted to clinical observations, including auxology, and the measurement of simple biochemical parameters (Table 1). Ob1, a female, had a normal birthweight of 3.46 kg, but gained weight rapidly in the early postnatal period. Her weight deviated from predicted centiles by 4 months of age (Fig. 3a). At the age of 8 years, Ob1 weighs 86 kg (>99.6th centile), her percentage body fat is 57% (reference range for children, 15–25%), and her height is 137 cm (75th centile). As a result of her obesity she developed abnormalities of growth in the long bones of her legs, resulting in the need for corrective limb surgery. She underwent liposuction of lower-limb fat in 1994 in an attempt to improve her mobility.

Ob2 is a cousin of Ob1 and is now aged 2 years. He had a normal weight at birth (3.53 kg) but rapidly became obese, deviating from predicted centiles for weight by 3 months of age (Fig. 3b). He currently weighs 29 kg (>99.6th centile), with 54% body fat. He has difficulty in walking because of extreme obesity. His current height is 89 cm (75th centile). In contrast to *ob/ob* mice, which show stunted linear growth, the heights of both Ob1 and Ob2 are at the 75th centile. However, given the age of the subjects it is not possible to comment on the future effects of leptin deficiency on the pubertal growth spurt and final adult height of these children.

Although a formal assessment of appetite has not been conducted, there is a clear history of marked hyperphagia, with both children noted from early infancy to be constantly hungry, demanding food continuously and eating considerably more than their siblings. Thus, in both mice and humans, congenital leptin deficiency is associated with a normal birthweight followed by the rapid development of severe obesity associated with hyperphagia and impaired satiety. Detailed assessment of energy expenditure has not yet been possible in these children, although their mean body temperatures are within the normal range (36–37 °C).

The *ob/ob* mice are also characterized by hypogonadotropic hypogonadism, resulting in sterility. As both Ob1 and Ob2 are both clinically prepubertal, with serum concentrations of luteinizing hormone, follicle-stimulating hormone, oestradiol and testosterone at prepubertal levels (Table 1), the effects of congenital leptin deficiency on the human reproductive system cannot yet be established in these children. In contrast to *ob/ob* mice, which are markedly hypercortisolaemic, plasma cortisol levels at 9:00 in both children are within the reference range, with Ob1 having a slight increase in 9:00 plasma adrenocorticotrophic hormone (Table 1). However, the administration of 1 mg of dexamethasone at 24:00 to Ob1 completely suppressed urinary free cortisol to <17 nmol l⁻¹.

Fasting plasma glucose was normal in both children, but fasting

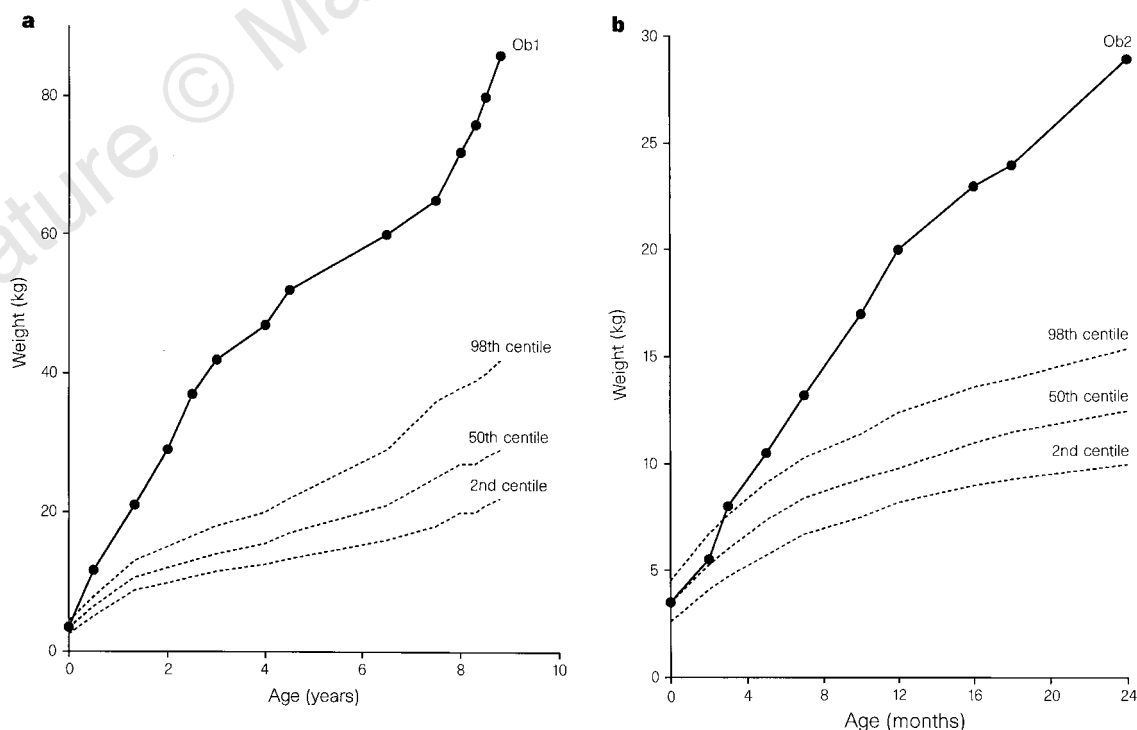


Figure 3 Ob1 and Ob2 had normal birthweights but rapidly developed severe postnatal obesity. **a**, Auxology of proband Ob1. Ob1's birthweight was at the 50th centile. By 4 months of age her weight was above the 98th centile. She is now aged 8 years. Her current height, 137 cm, is on the 75th centile for girls. Her current weight is 86 kg, and her mobility is severely impaired by her extreme obesity. **b**,

Auxology of proband Ob2. Ob2 is a first cousin of Ob1 and both sets of parents are related. Ob2's birthweight was on the 50th centile for boys. By 3 months his weight was above the 98th centile. He is now aged 2 years. His current height, 89 cm, is at the 75th centile. His current weight is 29 kg, and his mobility is also severely impaired by extreme obesity.

insulin levels were elevated in Ob1, consistent with the hyperinsulinaemia and insulin resistance seen in *ob/ob* mice (Table 1). Furthermore, the markedly higher plasma insulin concentration in Ob1 compared to Ob2 suggests that insulin resistance may worsen with age in human leptin-deficient states.

Although thyroid dysfunction is not a consistent feature of the phenotype of *ob/ob* mice¹⁵, a slight elevation of thyroid-stimulating hormone was noted in both children. The detection of a mild dysfunction of the pituitary–thyroid axis in these leptin-deficient children may represent a species difference, or may simply reflect the fact that normal ranges for serum thyroid-stimulating hormone have been defined more precisely for humans.

Thus, congenital deficiency of leptin in humans results in a phenotype with striking similarities to that seen in mice (severe obesity, hyperphagia and hyperinsulinaemia). The human probands seem to differ from *ob/ob* mice in that they do not show impaired linear growth, nor do they have marked hypercortisolaemia or hyperglycaemia. Whether the latter abnormalities will appear with time in the human probands, or are absent as a result of intrinsic species differences or the retention of a residual amount of leptin bioactivity in Ob1 and Ob2, remains to be established.

The fact that all four parents and one of the four siblings of the probands were heterozygous for the frame-shift mutation in the leptin gene allowed us to examine whether this genotype was associated with any phenotypic abnormalities in humans. None of the four heterozygous parents, nor the one heterozygous sibling, were morbidly obese (Fig. 2), a finding consistent with the absence of severe obesity in the murine *ob* heterozygotes¹⁶. Additionally, in the heterozygote relatives, the relationship between percentage body fat and serum leptin concentration was not significantly different from that seen in control subjects (Fig. 2). It should be pointed out that the control subjects in these studies were Caucasians from the UK, whereas the probands and their family members are all of Pakistani descent. However, to our knowledge, no consistent differences in serum leptin levels, other than those attributable to ethnic variations in body fat content, have been reported in subjects of different ethnic origins¹⁷. Thus, it is unlikely that the validity of the results is seriously affected by lack of matching for ethnicity. The absence of morbid obesity or a marked reduction in measurable serum leptin levels in the heterozygous relatives of the probands suggests either that fine tuning of leptin expression is not required for weight control in humans, or that compensatory mechanisms influencing expression of the wild-type allele are stimulated in the presence of an inactivating mutation in the other allele. However, definitive statements regarding potentially subtle effects of heterozygosity will require detailed studies of body composition and energy expenditure in these subjects and in ethnically matched controls.

In conclusion, the findings in subjects Ob1 and Ob2 strongly suggest that leptin critically influences energy balance in prepubertal humans. Our findings do not exclude an additional role for leptin in the initiation of puberty¹⁸ and/or the maintenance of gonadotropin function in postpubertal life¹⁹. Now that recombinant human leptin is available for investigation in humans, it should be possible to determine whether the correction of leptin deficiency in the two obese probands might have therapeutic benefits. □

Methods

Control subjects and clinical measurements. Adult control subjects were healthy Caucasians from the UK aged 30–39 years ($n = 30$). Prepubertal control subjects were Caucasian children from the UK aged 1–11 years that were undergoing elective surgery ($n = 16$). No control subject had undergone any recent marked change in body weight. Height (cm) and weight (kg) were measured, and percentage body fat for adults was calculated using Garrows formula²⁰, and estimated for children using Fomon's reference data²¹ with calculations for obese children based on the assumption that excess weight was 80% fat. The average percentage body fat for a 70-kg UK Caucasian male is

17%, and for a 60-kg UK Caucasian female is 22–25%. Average percentage body fat in prepubertal children ranges from 15% to 25% (ref. 21).

Serum leptin measurements. Serum samples were collected after an overnight fast and immediately frozen at -20°C . Samples were thawed once immediately before the assays. Radio-immunoassays were performed using a commercially available kit (Linco Research) with a detection limit of 0.5 ng ml^{-1} and in-house intra- and inter-assay coefficients of variation of 3% and 5%, respectively. Serum leptin levels of Ob1 and Ob2 were also measured using a highly sensitive enzyme-linked immunoassay (minimum detection limit, 70 pg ml^{-1} ; intra- and inter-assay coefficients of variation were 6.5% and 8.3%, respectively) (M. Nicolson, Amgen).

Cloning and sequencing of leptin cDNA. Total RNA was reverse-transcribed to first-strand cDNA using AMV reverse transcriptase (Promega). Leptin cDNA (nucleotides 11–527) was amplified by PCR and cloned into the pCR-Script SK(+) vector using the pCR Script cloning system (Stratagene). Four clones were sequenced using an ABI 373 automated sequencer.

Direct sequencing of leptin genomic DNA. Nucleotides 2536–2760 of leptin genomic DNA were amplified from genomic DNA of the probands and immediate family members by PCR. Sequencing of the PCR product was performed using an ABI 373 automated sequencer.

Single-strand conformation polymorphism analysis. Nucleotides 2536–2760 of the leptin gene were amplified from genomic DNA of probands and available family members by PCR including $[\alpha\text{-}^{32}\text{P}]\text{-dATP}$. The PCR product was diluted with formamide and electrophoresis performed on a 6% acrylamide gel at room temperature for 6–8 h at 20 W. The gel was dried and exposed to film for 24 h.

PCR primers and conditions. Amplification of full-length leptin cDNA of Ob2: sense primer, GAACCTGTGCGGATTCTTG; antisense primer, CAG-GAAGAGTGACCTCAAG; PCR, 1.5 mM MgCl_2 , (94°C 1 min, 55°C 1 min, 72°C 1 min) for 1 cycle, (94°C 1 min, 55°C 1 min, 72°C 1 min) for 40 cycles. Amplification of leptin from genomic DNA: sense primer, CCAACGACCTG-GAGAACCTCCGGGATC; antisense primer, CAGGAAGAGTGACCTCAAG; PCR, 1.5 mM MgCl_2 , (94°C 1 min, 55°C 1 min, 72°C 1 min) for 1 cycle, (94°C 1 min, 55°C 1 min, 72°C 1 min) for 30 cycles. Primer used in sequencing of PCR products: antisense, GTCCTGCAGAGACCCCTGCAGCCTGCT.

Transfection studies. Wild-type and mutant leptin cDNAs (Amgen) were inserted into the expression vector pRc. CMV (Invitrogen) using *Xba*I and *Apa*I. Sub-confluent CHO.K1 cells grown in four-well plates were transfected with the plasmids indicated using Lipofectamine according to the manufacturer's instructions (Gibco BRL). Approximately 24 h after transfection, immunoprecipitation of cell lysates and cultured media was performed using polyclonal anti-leptin antibody, PC3 (Amgen). Immunoprecipitates were electrophoresed on 15% SDS–PAGE, transferred to PVDF membrane (Immobilon), and probed with PC3 antibody and ^{125}I -labelled goat anti-rabbit secondary antibody (Cal Biochem). Results were analysed using a Fujix BAS 2000 phosphorimager.

Received 4 March; accepted 30 May 1997.

1. Zhang, Y. *et al.* Positional cloning of the mouse *obese* gene and its human homologue. *Nature* **372**, 425–432 (1994).
2. Maffei, M. *et al.* Increased expression in adipocytes of *ob* RNA in mice with lesions of the hypothalamus and with mutations at the *db* locus. *Proc. Natl Acad. Sci. USA* **92**, 6957–6960 (1995).
3. Stephens, T. W. *et al.* The role of neuropeptide Y in the antiobesity action of the *obese* gene product. *Nature* **377**, 530–534 (1995).
4. Vaisse, C. *et al.* Leptin activation of Stat3 in the hypothalamus of wild type and *ob/ob* mice but not *db/db* mice. *Nature Genet.* **14**, 95–97 (1996).
5. Pelleymounter, M. A. *et al.* Effects of the *obese* gene product on body weight regulation in *ob/ob* mice. *Science* **269**, 540–543 (1995).
6. Halaas, J. L. *et al.* Weight-reducing effects of the plasma protein encoded by the *obese* gene. *Science* **269**, 543–546 (1995).
7. Campfield, L. A., Smith, F. J., Guisez, Y., Devos, R. & Burn, P. Recombinant mouse *ob* protein: Evidence for a peripheral signal linking adiposity and central networks. *Science* **269**, 546–549 (1995).
8. Chen, G. *et al.* Disappearance of body fat in normal rats induced by adenovirus-mediated leptin gene therapy. *Proc. Natl Acad. Sci. USA* **93**, 14795–14799 (1996).
9. Maffei, M. *et al.* Absence of mutations in the human *ob* gene in obese/diabetic subjects. *Diabetes* **45**, 679–682 (1996).
10. Considine, R. V. *et al.* Evidence against either a premature stop codon or the absence of *obese* gene mRNA in human obesity. *J. Clin. Invest.* **95**, 2986–2988 (1995).
11. Bouchard, C. The causes of obesity: advances in molecular biology but stagnation on the genetic front. *Diabetologia* **39**, 1532–1533 (1996).
12. Maffei, M. *et al.* Leptin levels in human and rodent, measurement of plasma leptin and *ob* RNA in obese and weight-reduced subjects. *Nature Med.* **1**, 1155–1161 (1995).
13. Zhang, F. *et al.* Crystal structure of the *obese* protein leptin-E100. *Nature* **387**, 206–209 (1997).
14. Krawczak, M., Reiss, J. & Cooper, N. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* **90**, 41–54 (1992).

15. York, D. A., Otto, W. & Taylor, T. G. Thyroid status of obese (ob/ob) mice and its relationship to adipose tissue metabolism. *Comp. Biochem. Physiol.* **590**, 59–65 (1978).
16. Coleman, D. L. Obesity genes: beneficial effects in heterozygous mice. *Science* **203**, 663–665 (1979).
17. Ravussin, E. *et al.* Relatively low plasma leptin concentrations preceded weight gain in Pima Indians. *Nature Med.* **3**, 238–240 (1997).
18. Chehab, F. F., Mounzih, K., Lu, R. & Lim, M. E. Early onset of reproductive function in normal female mice treated with leptin. *Science* **275**, 88–90 (1997).
19. Ahima, R. S. *et al.* Role of leptin in the neuroendocrine response to fasting. *Nature* **382**, 250–252 (1996).
20. Garrow, J. S. & Webster, J. Quetelet's Index (W/H^2) as a measure of fatness. *Int. J. Obes.* **9**, 147–153 (1985).
21. Fomon, S. J., Haschke, F., Ziegler, E. E. & Nelson, S. E. Body composition of reference children from birth to age 10 years. *Am. J. Clin. Nutr.* **35**, 1169–1175 (1982).

Acknowledgements. We thank S. Jebb for help with analysis of body composition; C. N. Hales, S. Kumar, A. Kelly and P. Raggett for access to control samples; A. Roberts for help with clinical studies; M. McCaleb, F. Martin, N. Hernday, C. N. Hales, A. Krook, S. Jebb, A. Prentice and N. Finer for discussions; T. M. Cox, V. K. K. Chatterjee and P. Luzzio for reading the manuscript; and M. Flint for secretarial assistance. This research was supported by grants from the Wellcome Trust, the British Diabetic Association and a Medical Research Council (UK) studentship to C.T.M. H.R. is supported by the TMR programme of the EC.

Correspondence and requests for materials should be addressed to S.O.R. (e-mail: sorahill@hgmp.mrc.ac.uk).

Fringe modulates Notch–ligand interactions

Vladislav M. Panin*, Venizelos Papayannopoulos*, Richa Wilson* & Kenneth D. Irvine

Waksman Institute and Department of Molecular Biology and Biochemistry, Rutgers, The State University, Piscataway, New Jersey 08855, USA

* These authors contributed equally to this work

The Notch family of transmembrane receptor proteins mediate developmental cell-fate decisions¹, and mutations in mammalian Notch genes have been implicated in leukaemia, breast cancer, stroke and dementia^{2–4}. During wing development in *Drosophila*, the Notch receptor is activated along the border between dorsal and ventral cells^{5–7}, leading to the specification of specialized cells that express Wingless (Wg) and organize wing growth and patterning^{6,8,9}. Three genes, *fringe* (*fng*), *Serrate* (*Ser*) and *Delta* (*Dl*), are involved in the cellular interactions leading to Notch activation^{7,9–15}. *Ser* and *Dl* encode transmembrane ligands for Notch^{16,17}, whereas *fng* encodes a pioneer protein¹⁰. We have investigated the relationship between these genes by a combination of expression and coexpression studies in the *Drosophila* wing. We found that *Ser* and *Dl* maintain each other's expression by a positive feedback loop. *fng* is expressed specifically by dorsal cells and functions to position and restrict this feedback loop to the developing dorsal–ventral boundary. This is achieved by *fng* through a cell-autonomous mechanism that inhibits a cell's ability to respond to Serrate protein and potentiates its ability to respond to Delta protein.

Serrate protein (*Ser*) seems to act as a signal from dorsal cells that activates Notch in ventral cells, whereas Delta protein (*Dl*) seems to act as a signal from ventral cells that activates Notch in dorsal cells^{7,9,11–15}. Both of these proteins are broadly expressed during early wing development: *Dl* is initially expressed by all wing cells and *Ser* is initially expressed by all dorsal wing cells. However, their expression soon becomes restricted to the dorsal–ventral (D–V) boundary of the wing (Fig. 1b–d)^{7,9,11,13,14}. This observation suggests that, as well as being involved in signalling between dorsal and ventral cells, *Ser* and *Dl* also respond to D–V cell interactions, and so they may regulate each other's expression. To examine this, transgenic flies containing *UAS-Ser* (ref. 12) or *UAS-Dl* (ref. 14) were crossed to flies containing *ptc-Gal4* (ref. 18), which drives expression in a stripe along the anterior side of the anterior–posterior (A–P) compartment boundary (Fig. 1a)¹⁹. The ability of *Ser* to induce *Dl* expression, and of *Dl* to induce *Ser* expression, was then assayed by antibody staining of mid third-instar wing imaginal discs.

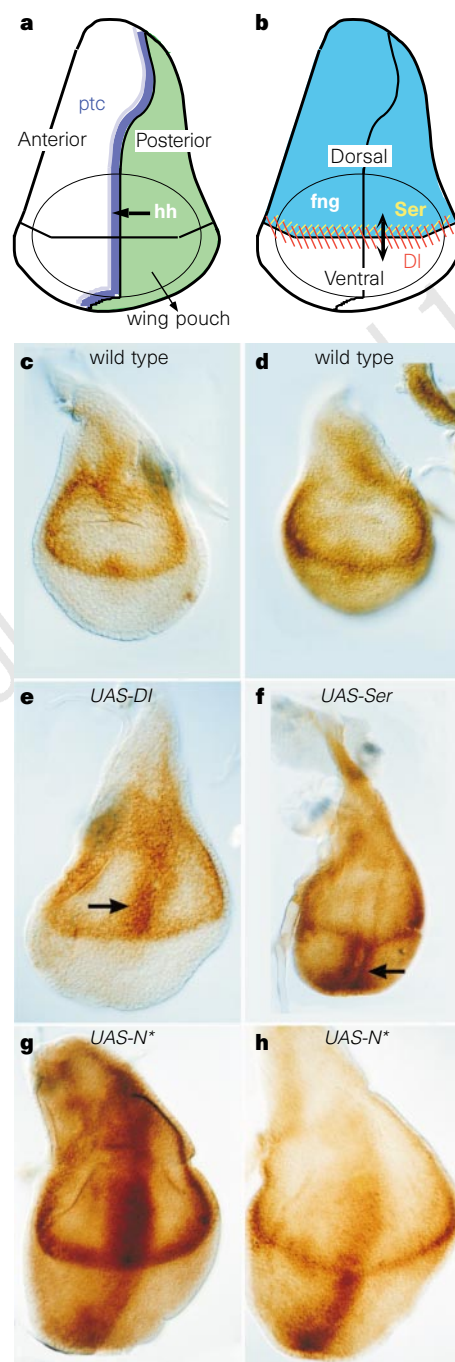


Figure 1 *Ser* and *Dl* expression and feedback regulation in the wing imaginal disc. Dorsal is up, anterior to the left. **a**, Schematic representation of A–P compartmental division of the wing disc; expression of the P-compartment signal, Hedgehog (Hh, green), and of Patched (Ptc, purple) is shown. The wing pouch, which is the portion of the disc that gives rise to the adult wing blade, is demarcated by an oval. **b**, Schematic representation of D–V compartmental division of the wing disc. The expression patterns of *Fng* (blue), *Ser* (yellow hatching) and *Dl* (red hatching) are dynamic; expression from early to mid-third instar is shown. **c–h**, Expression in mid third instar wing imaginal discs (at same magnification) of *Ser* (**c**, **e**, **g**) and *Dl* (**d**, **f**, **h**). **c**, **d**, Wild type. **e**, *UAS-Dl30A ptcGal4* raised at 25 °C; arrow points to ectopic expression of *Ser* in dorsal cells. **f**, *UAS-Ser ptcGal4* raised at 25 °C; arrow points to ectopic expression of *Dl* in ventral cells. Previous studies of *Ser* signalling have indicated that cells expressing *Ser* are themselves refractory to *Ser*; this presumably accounts for the restriction of *Dl* induction to the edges of the *Ptc* expression stripe. **g**, **h**, *UAS-activated-Notch ptc-Gal4* raised at 18 °C.

These experiments confirmed that Ser and Dl induce each other's expression (Fig. 1e,f). Importantly, their ability to do this is dorsal–ventrally asymmetric, because Ser induces Dl strongly in ventral cells, but only very weakly in dorsal cells, whereas Dl induces Ser expression in dorsal cells, but not in ventral cells. This reciprocal induction of Ser by Dl and of Dl by Ser places them in a positive feedback loop. However, during normal wing development this feedback loop must be restricted to the developing D–V border; if this were not the case, then Ser and Dl expression, and consequently Notch activation, would spread throughout the wing. The feedback loop is restricted to the D–V border by the complementary restrictions on the ability of dorsal and ventral cells to respond to Ser and Dl.

Three observations suggested that *fng* could bring about this differential responsiveness of dorsal and ventral cells to Notch ligands. First, *fng* is expressed and required specifically in dorsal cells for wing formation (Fig. 1b)¹⁰. Second, changes in the *fng* expression pattern can lead to ectopic wing margin formation and alterations in wing growth^{10,11}. Third, *fng* affects the expression of

Ser¹¹ and Dl (Fig. 2c,d), which, because of the Ser–Dl feedback loop, could derive from an influence on their signalling activity.

To determine the effects of Fringe protein (Fng) on Ser and Dl activity, we assayed margin gene expression and margin bristle formation while coexpressing these proteins in ventral cells. Fng was misexpressed by taking advantage of a dominant allele, *fng*^{D4}, that is associated with ubiquitous low-level misexpression of Fng in ventral cells¹⁰. Ser and Dl were then expressed under UAS control by using the *ptc-Gal4* driver. Importantly, because Ser and Dl are expressed under the control of *patched* (*ptc*), rather than under their own promoters, any effects of Fng should reflect its influence on their activity rather than on their transcription.

Ectopic expression of Ser under *ptc-Gal4* control leads to ectopic wing-margin gene expression and adult wing-margin formation in ventral cells along the edges of the ectopic Ser stripe (Figs 1f, 2b, 3c, 4b)^{9,11–13}. Misexpression of Fng in ventral cells inhibits these effects of Ser activity (Fig. 3g,h, 4e), demonstrating that Fng can inhibit Ser signalling. The inhibition is incomplete, presumably because ventral *fng* expression in *fng*^{D4} animals is low relative to normal dorsal levels of *fng*¹⁰. Complementary experiments with a hypomorphic *fng* allele demonstrate that the ability of Ser to signal to dorsal cells is increased when Fng activity is reduced (Fig. 2a, and data not shown).

Misexpression of Dl under *ptc-Gal4* control induces ectopic wing-margin formation and Ser expression in dorsal cells but not in ventral cells¹⁴ (Figs 1e, 3a,b, 4e). However, when Fng is misexpressed in ventral cells as a consequence of the *fng*^{D4} mutation, Dl induces both Ser expression and margin formation in both dorsal and ventral cells (Figs 3e,f, 4f). Thus Fng potentiates Dl signalling, allowing ventral cells to respond to Dl just as dorsal cells normally do.

These results demonstrate that Fng exerts opposing influences on the activity of the two known Notch ligands in *Drosophila*, inhibiting Ser and potentiating Dl. The experiments described below investigate whether these effects of Fng are exerted on a cell's ability to send or receive ligand signals, and whether Fng exerts its effects solely on cells expressing Fng (that is, cell autonomously) or also on neighbouring cells. These questions were addressed by examining the consequences of coincident, as opposed to ubiquitous, misexpression of Fng with Notch ligands.

Coexpression of Fng with Ser in *UAS-fng UAS-Ser ptc-Gal4* flies inhibits Ser signalling along the anterior side of the *ptc* expression stripe, consistent with the results of coexpression in *fng*^{D4} animals (Figs 2e,f, 3o,p, 4h). In contrast, however, Ser signalling was not inhibited along the posterior side of the *ptc* expression stripe. *ptc* expression declines in a gradient away from the A–P compartment boundary, and Ser induces Wg and Dl expression along the anterior side of the stripe in cells that actually have low levels of *ptc*-driven expression (Figs 2b, 3d). In contrast, there is a sharp juxtaposition between *ptc*-expressing and non-expressing cells along the posterior side of the stripe, and Ser induces Wg and Dl expression in adjacent, non-*ptc*-expressing cells^{9,11} (Figs 2b, 3d). Thus, when *fng* and Ser are coexpressed under *ptc-Gal4* control, the cells that would have formed the anterior ectopic margin now express Fng, whereas those that would have formed the posterior ectopic margin do not (Figs 3d, 4e inset). The distinct effects of Fng on Ser-induced margin formation along the anterior versus the posterior sides of the *ptc* expression stripe therefore indicate that Fng inhibits Ser activity only when it is expressed in receiving cells, and not when its expression is restricted to Ser-signalling cells. This conclusion is also consistent with the natural expression patterns of *fng* and Ser: they are coincidentally expressed along the dorsal side of the D–V compartment boundary^{9–11,13}.

Fng was also coincidentally expressed with Dl using UAS and *ptc-Gal4* transgenes. In the adult wing, this results in the induction of a band of wing-margin tissue in ventral cells (Fig. 4i). In the third-instar imaginal disc, some Ser expression is induced by Fng alone,

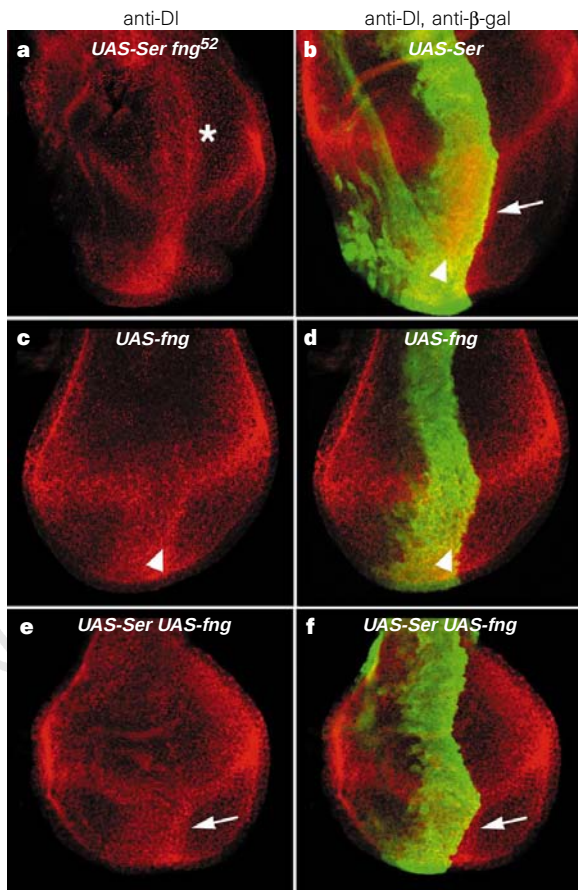


Figure 2 Dl expression induced by Fng and Ser. Third-instar wing imaginal discs, with ventral down, anterior to the left. Dl expression is shown in red; β -galactosidase, which marks cells expressing *ptc-Gal4*, is shown in green. Arrows point to Dl expression outside the *ptc-Gal4* stripe; arrowheads point to expression within the stripe. **a**, *fng*⁵² *UAS-Ser ptc-Gal4* raised at 25°C. Dl expression is now induced in dorsal cells (asterisk) at levels similar to endogenous Dl along the D–V boundary (see Fig. 1f). **b**, *UAS-Ser UAS-lacZ ptc-Gal4* raised at 25°C. **c**, *UAS-fng UAS-lacZ ptc-Gal4* raised at 25°C. Expression of *fng* itself is not affected by Dl (not shown). **d**, *UAS-fng UAS-lacZ ptc-Gal4* raised at 25°C. **e**, *UAS-Ser UAS-fng UAS-lacZ ptc-Gal4* raised at 25°C. As in *UAS-fng ptc-Gal4*, but in contrast to *UAS-Ser-ptc-Gal4* (see **b**, and Fig. 1f), Dl staining appears weak and diffuse along the anterior of the *Ptc* expression stripe.

although this induction is weak relative to endogenous Ser levels (Fig. 3i)¹¹. When Fng and DI are coincidentally expressed under *ptc-Gal4* control, Ser expression is induced at high levels within the *ptc* expression stripe (Fig. 3m,n). However, no Ser expression is induced in adjacent ventral cells. This contrasts with the induction of Ser expression by DI in adjacent Fng-expressing cells (Fig. 3b,f,n). These observations demonstrate that Fng potentiates DI signalling, but only when it is expressed by cells receiving the DI signal. This is also consistent with their natural expression patterns: DI expression is required mainly in ventral cells, whereas only dorsal, Fng-expressing cells respond significantly to DI^{7,14}.

The Fng protein enables cells in the wing to respond to DI by activating Ser. However, the distinct expression patterns of Ser and DI in the embryo indicate that DI does not generally induce Ser expression^{20,21}. This raises the question of whether Fng exerts a

qualitative effect on DI signalling and actually changes the way that cells respond to the interaction between DI and Notch. Alternatively, the effect of Fng could simply be quantitative. For example, Ser expression may be induced only at relatively high levels of Notch activation. To help distinguish between these possibilities, we examined Ser expression in wings expressing a ligand-independent activated Notch protein^{14,22}. These experiments showed that ventral wing cells, which lack Fng, have the capacity to respond to Notch activation by expressing Ser (Fig. 1g). Constitutive activation of Notch also induces DI expression in both dorsal and ventral wing cells (Fig. 1h). The difference in levels of Ser and DI expression between dorsal and ventral cells in these experiments is consistent with the observations that expression of both of these proteins is induced by activated Notch, and that they positively regulate each other's expression in a spatially restricted manner (Fig. 1e,f). The

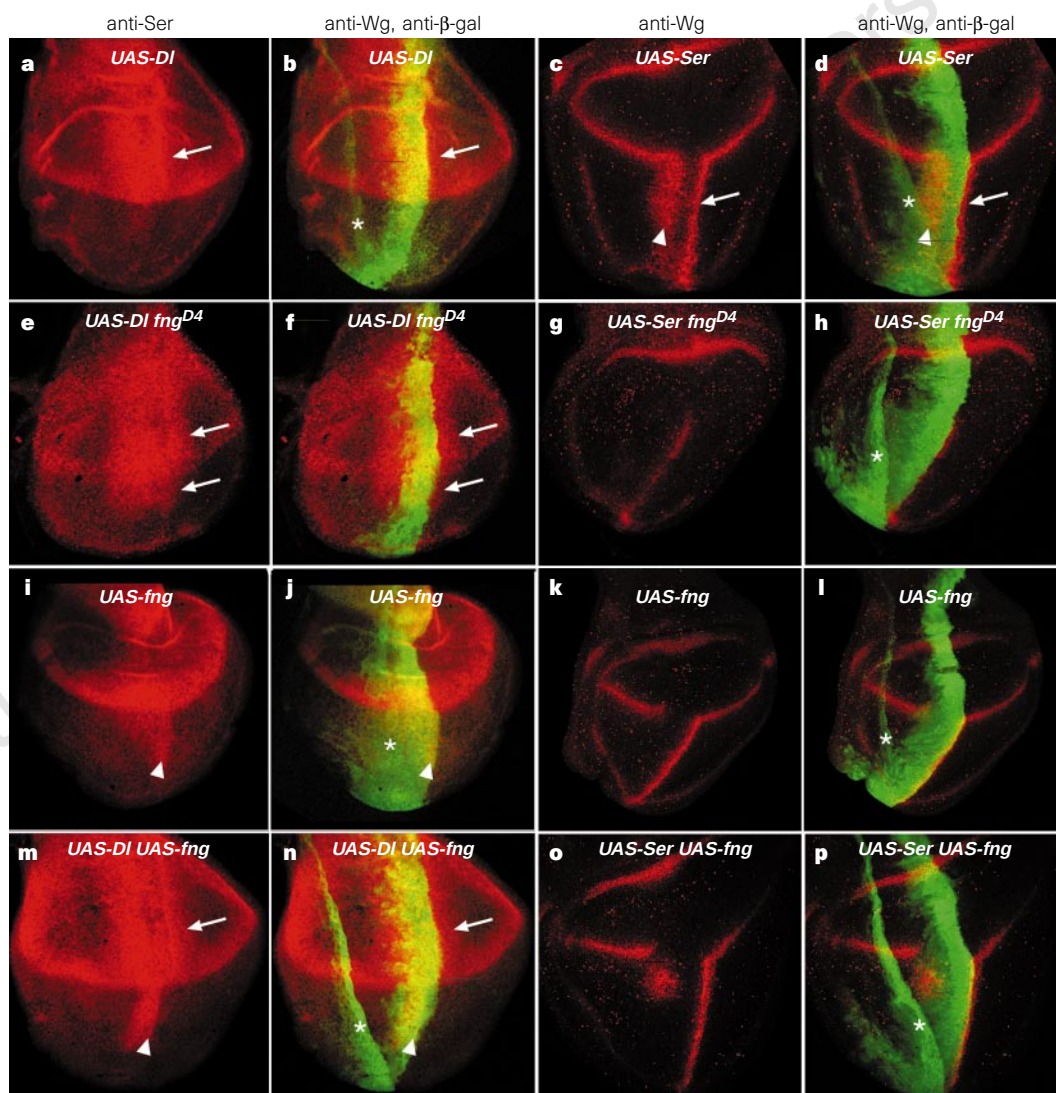


Figure 3 Effects of Fng on Ser and DI activity in the wing imaginal disc. Third-instar wing imaginal discs, with ventral down, anterior to the left. Ser (a, b, e, f, i, j, m, n) and Wg (c, d, g, h, k, l, o, p) expression are shown in red; β -galactosidase, which marks *ptc-Gal4* expressing cells, is shown in green. Arrows point to Ser and Wg expression outside the *ptc-Gal4* stripe; arrowheads point to expression within the stripe; and asterisks mark the *ptc-Gal4* stripe in the peripodial membrane. a, b, *UAS-DI130B UAS-lacZ ptc-Gal4* raised at 18 °C. c, d, *UAS-Ser UAS-lacZ ptc-Gal4* raised at 25 °C. e, f, *fngD4 UAS-DI130B UAS-lacZ ptc-Gal4*, raised at 18 °C. Ubiquitous expression of Fng reduces normal Ser margin expression, but enables DI to induce ectopic Ser expression in ventral cells (bottom arrow). g, h,

fngD4 UAS-Ser UAS-lacZ ptc-Gal4, raised at 25 °C. i–l, *UAS-fng UAS-lacZ ptc-Gal4*, raised at 25 °C. We have re-examined the induction of Ser expression by Fng and, in contrast to a previous report¹¹, we find that Ser is only induced within Fng-expressing cells. m, n, *UAS-fng UAS-DI130B UAS-lacZ ptc-Gal4* raised at 18 °C. The ventral Ser expression stripe is actually narrower than the Ptc expression stripe. The failure of the most posterior cells within the stripe, which have the highest levels of DI and Fng, to induce Ser expression suggests that DI signalling is sensitive to the relative levels of Fng, DI and Notch. o, p, *UAS-fng UAS-Ser UAS-lacZ ptc-Gal4* raised at 25 °C.

similar effects of activated Notch on dorsal and ventral cells also imply that Fng exerts its effects upstream of Notch activation^{9,14,23}.

Although the amino-acid sequence of Fng predicts that it is a secreted protein¹⁰, our results indicate that Fng functions cell-autonomously to modulate Notch signalling. To investigate the distribution of Fng protein, we constructed epitope-tagged proteins and expressed them both in cultured *Drosophila* cells and in imaginal discs. In the wing, epitope-tagged Fng generates phenotypes indistinguishable from those of wild-type Fng, indicating that it is functional (not shown). Strikingly, immunostaining of both cultured cells and imaginal discs revealed that some Fng protein localizes to the periphery of Fng-expressing cells (Fig. 5). Some Fng protein is also secreted in diffusible form, although co-transfection (not shown) and cell-mixing experiments (Fig. 5a) confirmed that plasma-membrane staining of Fng is only observed on cells expressing

Fng.

These localization studies, together with the observations that Fng cell-autonomously affects a cell's ability to respond to Notch ligands and that activated Notch proteins seem to be Fng-independent, imply that the activity of cell-associated Fng protein differentially modulates the binding and/or activation of Notch by its two known *Drosophila* ligands. Although Fng is not required for all Notch-mediated cell-fate decisions in *Drosophila* (K.D.I., unpublished data), our results, as well as those of others^{24–28}, suggest that the functional relationship between Fng and Notch signalling we have defined is likely to be essential for the development of a range of metazoan tissues.

The specification of morphogen-secreting cells along the borders between distinct cell populations has been proposed as a general developmental mechanism^{29,30}. In the *Drosophila* wing, the com-

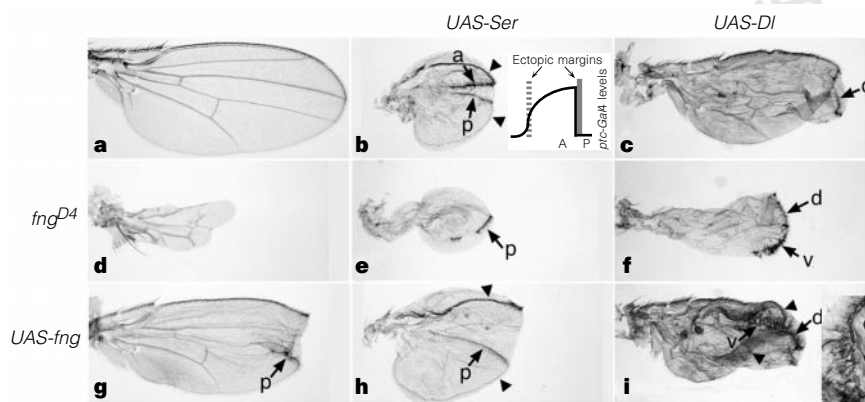


Figure 4 Wing phenotypes associated with Fng, Ser and DI misexpression. Adult wings with proximal left and anterior up, except in **f**, in which dorsal is up; all wings are shown at the same magnification. Arrows point to anterior (a), posterior (p), dorsal (d) and ventral (v) ectopic margin. Arrowheads point to normal margin. **a**, Wild type. **b**, *UAS-Ser ptc-Gal4* raised at 25°C. Inset, a schematic representation showing that *ptc*-driven expression (black line) is absent in the posterior compartment and graded away from the A-P boundary. The posterior ectopic margin (continuous shaded line) is composed of posterior-row hairs and forms in cells that lack Ser; the anterior ectopic margin (dashed line) is composed of double-row bristles and forms in cells

with low-level Ser expression. This anterior ectopic margin does not form in the presence of Fng (**e**, **h**). **c**, *UAS-DI30A ptc-Gal4* raised at 28°C. **d**, *fng^{D4}*. **e**, *fng^{D4} UAS-Ser ptc-Gal4* raised at 25°C. **f**, *fng^{D4} UAS-DI30A ptc-Gal4* raised at 28°C. **g**, *UAS-fng ptcGal4*. **h**, *UAS-Ser UAS-fng ptcGal4* raised at 25°C. Margin formation seems to be stimulated by coexpression of *fng* with *Ser*, as the posterior ectopic margin here extends to the base of wing. **i**, *UAS-DI30A UAS-fng ptcGal4* raised at 28°C. Inset, close-up of the ectopic margin showing that the broad band of margin tissue that forms in ventral cells is clearly distinct from the row of margin bristles induced by *fng* alone (**g**).

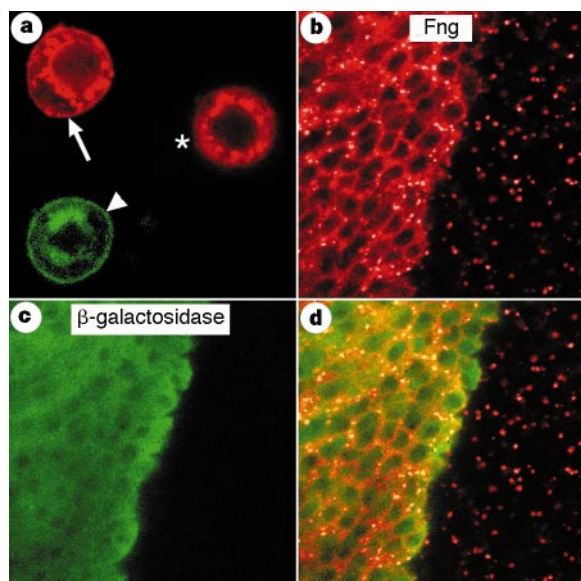


Figure 5 Immunolocalization of Fng protein. **a**, S2 cells derived from a mixture of two independently transfected cell populations, one expressing Fng-HA (red), the other expressing Notch (green). The Fng protein can be detected on the surface of many *fng*-expressing cells (arrow), although others lack detectable membrane staining (asterisk). The Fng protein is also secreted into the media, but this diffusible Fng does not detectably bind to other cells, indicated here by the absence of Fng staining on the Notch-expressing cell (arrowhead). **b–d**, Close-up of a *UAS-fngHA UAS-lacZ ptc-Gal4* third-instar wing disc, centred on the AP compartment boundary. **b**, Fng-HA (red). **c**, β -galactosidase (green). **d**, Superposition of fng-HA and β -galactosidase. Bright dots of Fng staining are detected both inside and outside the *ptc-Gal4* expression stripe, indicative of diffusible protein. Significant Fng-HA staining is also detected around the periphery of cells within the *ptc* expression stripe. This staining pattern is clearly distinct from the cytoplasmic staining of β -galactosidase.

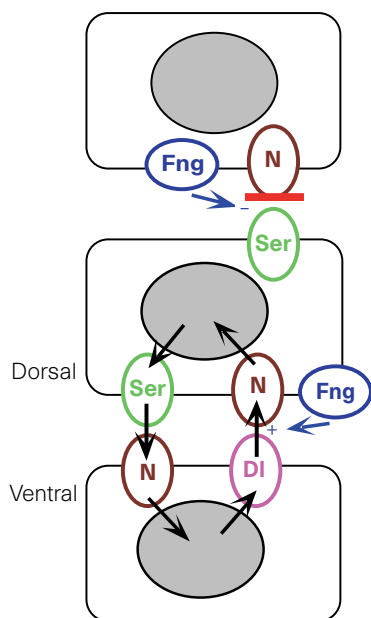


Figure 6 Schematic representation of signalling interactions at the dorsal-ventral compartment border. DI and Notch (N) are broadly expressed during early wing development, while Fng and Ser are restricted to dorsal cells. For clarity, not every protein or potential protein interaction is indicated. Expression of Fng allows dorsal cells to respond to DI, resulting in Notch activation. This leads to transcription of downstream genes, including Ser. Ser signals back from dorsal to ventral cells, activating Notch, and leading to the transcription of downstream genes, including DI. Fng blocks the ability of Ser to signal to other dorsal cells. Similarly, ectopic expression of Fng and Ser (for example, *UAS-fng UAS-Ser ptc-Gal4*) can establish ectopic Ser-DI feedback loops centred on the Fng expression border.

plementary effects of Fng on a cell's responsiveness to Ser and DI, its expression specifically by dorsal cells, and the Ser-DI feedback loop together help to position and restrict Notch activation to cells along the D-V compartment border (Fig. 6). This mechanism leads to the activation of Notch, and consequently expression of Wg, symmetrically in cells on both sides of the compartment border. In contrast, the A-P signalling system is unidirectional and leads to expression of the Decapentaplegic morphogen solely along the anterior side of the compartment boundary²⁹. These differences seem to have fundamental morphological consequences, as the anterior compartment of the wing is larger than the posterior compartment, whereas the dorsal and ventral compartments are exactly equal in size, and must be so for the wing to flatten and form correctly. □

Methods

Immunohistochemistry. For staining of wing imaginal discs, staged larvae were dissected in *Drosophila* Ringers at 4 °C and fixed in 4% paraformaldehyde, 0.1 M MOPS for 10 min at 25 °C. Tissue culture cells were collected and fixed 18–22 h after induction as described¹⁶. Antibody stains were performed using as primary antibodies mouse anti-DI 9B at 1:1,000, rabbit anti-Ser at 1:40, rabbit anti-Wg at 1:500, mouse anti-haemagglutinin epitope (HA) at 1:7,500 (Babco), rat-1 anti-Notch at 1:1,000, rabbit anti-β-galactosidase at 1:4,000 (Cappel), and mouse anti-β-galactosidase at 1:2,000 (Sigma). Histochemical detection was performed using the ABC system (Vector), and immunofluorescent detection was performed using preabsorbed FITC and Cy3 conjugated donkey anti-IgGs (Jackson ImmunoResearch). After histochemical stains, discs were dissected and mounted in 90% glycerol. After immunofluorescent stains, discs or cultured cells were mounted in Vectashield (Vector). Confocal images were obtained using a Molecular Dynamics MultiProbe 2001.

Expression constructs, cell culture and transfection. Double-stranded oligonucleotides encoding three copies of the haemagglutinin epitope (YPYDVPDYA) were cloned into a *Bgl*II site introduced after the last *fng* codon. The resulting construct was verified by DNA sequencing and subcloned into the pMHy vector³¹ for cell-culture expression and into the pUAST vector¹⁹ for *Drosophila* transformation. Expression plasmids were transfected into cultured cells by lipofection; a detailed protocol is available from the authors on request.

Received 3 March; accepted 30 May 1997.

- Artavanis-Tsakonas, S., Matsuno, K. & Fortini, M. Notch signaling. *Science* **268**, 225–230 (1995).
- Joutel, A. *et al.* *Notch3* mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383**, 707–710 (1996).
- Robbins, J., Blondel, B. J., Gallahan, D. & Callahan, R. Mouse mammary tumor gene *int-3*: a member of the *Notch* gene family transforms mammary epithelial cells. *J. Virol.* **66**, 2594–2599 (1992).
- Ellisen, L. W. *et al.* *TAN-1*, the human homolog of the *Drosophila Notch* gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell* **66**, 649–661 (1991).
- de Celis, J. F. & Garcia-Bellido, A. Roles of the *Notch* gene in *Drosophila* wing morphogenesis. *Mech. Dev.* **46**, 109–122 (1994).
- Rulifson, E. J. & Blair, S. S. *Notch* regulates *wingless* expression and is not required for reception of the paracrine wingless signal during wing margin neurogenesis in *Drosophila*. *Development* **121**, 2813–2824 (1995).
- de Celis, J. F., Garcia-Bellido, A. & Bray, S. J. Activation and function of Notch at the dorsal-ventral boundary of the wing imaginal disc. *Development* **122**, 359–369 (1996).
- Zecca, M., Basler, K. & Struhl, G. Direct and long range action of a wingless morphogen gradient. *Cell* **87**, 833–844 (1996).
- Diaz-Benjumea, F. J. & Cohen, S. M. Serrate signals through Notch to establish a Wingless-dependent organizer at the dorsal/ventral compartment boundary of the *Drosophila* wing. *Development* **121**, 4215–4225 (1995).
- Irvine, K. D. & Wieschaus, E. *fringe*, a boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during *Drosophila* wing development. *Cell* **79**, 595–606 (1994).
- Kim, J., Irvine, K. D. & Carroll, S. B. Cell interactions and inductive signals at the dorsal/ventral boundary of the developing *Drosophila* wing. *Cell* **82**, 795–802 (1995).
- Speicher, S. A., Thomas, U., Hinz, U. & Knust, E. The *Serrate* locus of *Drosophila* and its role in morphogenesis of the wing imaginal discs: control of cell proliferation. *Development* **120**, 535–544 (1994).
- Couso, J. P., Knust, E. & Martinez Arias, S. *Serrate* and *wingless* cooperate to induce *vestigial* gene expression and wing formation in *Drosophila*. *Curr. Biol.* **5**, 1437–1448 (1995).
- Doherty, D., Feger, G., Younger-Shepherd, S., Jan, L. Y. & Jan, Y. N. Delta is a ventral to dorsal signal complementary to Serrate, another Notch ligand, in *Drosophila* wing formation. *Genes Dev.* **10**, 421–434 (1996).
- Jönsson, F. & Knust, E. Distinct functions of the *Drosophila* genes *Serrate* and *Delta* revealed by ectopic expression during wing development. *Dev. Genes Evol.* **206**, 91–101 (1996).
- Fehon, R. G. *et al.* Molecular interactions between the protein products of the neurogenic loci *Notch* and *Delta*, two EGF-homologous genes in *Drosophila*. *Cell* **61**, 523–534 (1990).
- Rebay, L. *et al.* Specific EGF repeats of Notch mediate interactions with Delta and Serrate: Implications for Notch as a multifunctional receptor. *Cell* **67**, 687–699 (1991).
- Hinz, U., Giebel, B. & Campos-Ortega, J. The basic helix-loop-helix domain of *Drosophila* lethal scute protein is sufficient for proneural function and activates neurogenic genes. *Cell* **76**, 77–87 (1994).
- Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Thomas, U., Speicher, S. A. & Knust, E. The *Drosophila* gene *Serrate* encodes an EGF-like transmembrane protein with a complex expression pattern in embryos and wing discs. *Development* **111**, 749–761 (1991).
- Kopczynski, C. C. & Muskavitch, M. A. T. Complex spatio-temporal accumulation of alternative transcripts from the neurogenic gene *Delta* during embryogenesis. *Development* **107**, 623–636 (1989).
- Struhl, G., Fitzgerald, K. & Greenwald, I. Intrinsic activity of the Lin-12 and Notch intracellular domains in vivo. *Cell* **74**, 331–345 (1993).
- Kim, J. *et al.* Integration of positional signals and regulation of wing formation and identity by *Drosophila* *vestigial* gene. *Nature* **382**, 133–138 (1996).
- Fleming, R. J., Gu, Y. & Hukriede, N. A. *Serrate*-mediated activation of *Notch* is specifically blocked by the product of the gene *fringe* in the dorsal compartment of the *Drosophila* wing imaginal disc. *Development* (in the press).
- Wu, J. Y., Wen, L., Zhang, W.-J. & Rao, Y. The secreted product of *Xenopus* gene *lunatic Fringe*, a vertebrate signaling molecule. *Science* **273**, 355–358 (1996).
- Johnston, S. H. *et al.* A family of mammalian *Fringe* genes implicated in boundary determination and the *Notch* pathway. *Development* **124**, 2245–2254 (1997).
- Laufer, E. *et al.* Expression of *Radical fringe* in limb-bud ectoderm regulates apical ectodermal ridge formation. *Nature* **386**, 366–373 (1997).
- Rodriguez-Esteban, C. *et al.* *Radical fringe* positions the apical ectodermal ridge at the dorsoventral boundary of the vertebrate limb. *Nature* **386**, 360–366 (1997).
- Lawrence, P. A. & Struhl, G. Morphogens, compartments, and pattern: Lessons from *Drosophila*? *Cell* **85**, 951–961 (1996).
- Meinhardt, H. Cell determination boundaries as organizing regions for secondary embryonic fields. *Dev. Biol.* **96**, 375–385 (1983).
- Koelle, M. R. *et al.* The *Drosophila* *EcR* gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell* **67**, 59–77 (1991).

Acknowledgements. We thank E. Knust, D. Doherty, S. Shepard, S. Artavanis-Tsakonas, R. Flemming, M. Muskavitch, R. Nusse, S. Carroll, N. Perrimon, L. Cherbas, C. Thummel, E. Parker and the Bloomington *Drosophila* Stock Center for *Drosophila* stocks, antibodies, plasmids, cultured cell lines, and advice; R. Flemming for communicating results before publication; and C. Rauskolb, E. Wieschaus, T. Vogt, R. Padgett, G. Struhl and R. Steward for comments on the manuscript. This work was supported by an ACS junior faculty research award from the Cancer Institute of New Jersey, a grant from the Charles and Joanna Busch Memorial Fund at Rutgers, The State University, and a grant from the NIGMS, NIH to K.D.I.

Correspondence and requests for materials should be addressed to K.D.I. (e-mail: irvine@mbcl.rutgers.edu).

Robustness in simple biochemical networks

N. Barkai & S. Leibler

Departments of Physics and Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA

Cells use complex networks of interacting molecular components to transfer and process information. These "computational devices of living cells"¹ are responsible for many important cellular processes, including cell-cycle regulation and signal transduction. Here we address the issue of the sensitivity of the networks to variations in their biochemical parameters. We propose a mechanism for robust adaptation in simple signal transduction networks. We show that this mechanism applies in particular to bacterial chemotaxis²⁻⁷. This is demonstrated within a quantitative model which explains, in a unified way, many aspects of chemotaxis, including proper responses to chemical gradients⁸⁻¹². The adaptation property^{10,13-16} is a consequence of the network's connectivity and does not require the 'fine-tuning' of parameters. We argue that the key properties of biochemical networks should be robust in order to ensure their proper functioning.

Cellular biochemical networks are highly interconnected: a perturbation in reaction rates or molecular concentrations may affect numerous cellular processes. The complexity of biochemical networks raises the question of the stability of their functioning. One possibility is that to achieve an appropriate function, the reaction rate constants and the enzymatic concentrations of a network need to be chosen in a very precise manner, and any deviation from the 'fine-tuned' values will ruin the network's performance. Another

possibility is that the key properties of biochemical networks are robust; that is, they are relatively insensitive to the precise values of biochemical parameters. Here we explore the issue of robustness of one of the simplest and best-known signal transduction networks: a biochemical network responsible for bacterial chemotaxis. Bacteria such as *Escherichia coli* are able to sense (temporal) gradients of chemical ligands in their vicinity². The movement of a swimming bacterium is composed of a series of 'smooth runs', interrupted by events of 'tumbling', in which a new direction for the next run is chosen randomly. By modifying the tumbling frequency, a bacterium is able to direct its motion either towards attractants or away from repellents. A well established feature of chemotaxis is its property of adaptation^{10,13-16}: the steady-state tumbling frequency in a homogeneous ligand environment is insensitive to the value of ligand concentration. This property allows bacteria to maintain their sensitivity to chemical gradients over a wide range of attractant or repellent concentrations.

The different proteins that are involved in chemotactic response have been characterized in great detail, and much is known about the interactions between them (Fig. 1a). In particular, the receptors that sense chemotactic ligands are reversibly methylated. Biochemical data indicate that methylation is responsible for the adaptation property: changes in methylation of the receptor can compensate for the effect of ligand on tumbling frequency. Theoretical models proposed in the past assumed that the biochemical parameters are fine-tuned to preserve the same steady-state behaviour at different ligand concentrations^{17,18}. We present an alternative picture in which adaptation is a robust property of the chemotaxis network and does not rely on the fine-tuning of parameters.

We have analysed a simple two-state model of the chemotaxis network closely related to the one proposed previously^{2,19}. The two-state model assumes that the receptor complex has two functional states: active and inactive. The active receptor complex shows a kinase activity: it phosphorylates the response regulator molecules,

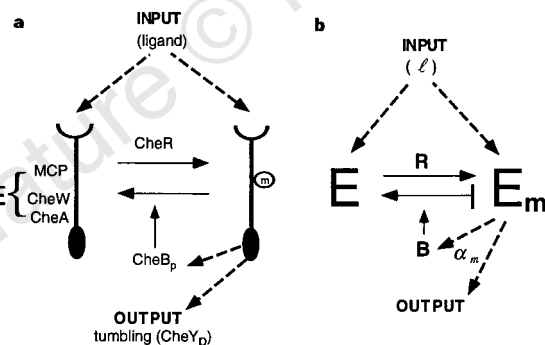


Figure 1 a, The chemotaxis network. Chemotactic ligands bind to specialized receptors (MCP) which form stable complexes (E), with the proteins CheA and CheW. CheA is a kinase that phosphorylates the response regulator, CheY, whose phosphorylated form (CheY_p) binds to the flagellar motor and generates tumbling. Binding of the ligand to the receptor modifies the tumbling frequency by changing the kinase activity of CheA. The receptor can also be reversibly methylated. Methylation enhances the kinases activity and mediates adaptation to changes in ligand concentration. Two proteins are involved in the adaptation process: CheR methylates the receptor, CheB demethylates it. A feedback mechanism is achieved through the CheA-mediated phosphorylation of CheB, which enhances its demethylation activity. **b**, Mechanism for robust adaptation. E is transformed to a modified form, E_m, by the enzyme R; enzyme B catalyses the reverse modification reaction. E_m is active with a probability of $\alpha_m(I)$, which depends on the input level I . Robust adaptation is achieved when R works at saturation and B acts only on the active form of E_m. Note that the rate of reverse modification is determined by the system's output and does not depend directly on the concentration of E_m (vertical bar at the end of the arrow).

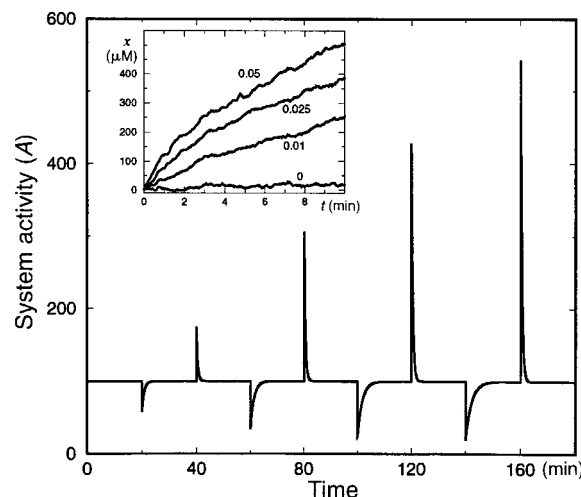


Figure 2 Chemotactic response and adaptation. The system activity, A , of a model system (the reference system described in Methods) which was subject to a series of step-like changes in the attractant concentration, is plotted as a function of time. Attractant was repeatedly added to the system and removed after 20 min, with successive concentration steps of I of 1, 3, 5 and 7 μM . Note the asymmetry to addition compared with removal of ligand, both in the response magnitude and the adaptation time. The chemotactic drift velocity of this system is presented in the inset. Inset: the different curves correspond to gradients $\nabla I = 0, 0.01, 0.025$ and $0.05 \mu\text{M}/\mu\text{m}$. An average change in receptor occupancy of less than 1% per second is sufficient to induce a mean drift velocity of the order of microns per second.

which then bind to the motors and induce tumbling. The receptor complexes can be either in the active or in the inactive state, although with probabilities that depend on both their methylation level and ligand occupancy. The average complex activity can be considered as the output of the network, whereas its input is the concentration of the ligand. A quantitative description of the model consists of a set of coupled differential equations describing interactions between protein components (Box 1).

The two-state model correctly reproduces the main features of bacterial chemotaxis. When a typical model system is subject to a step-like change in attractant concentration, l (Fig. 2), it is able to respond and to adapt to the imposed change. The adaptation is nearly perfect for all ligand concentrations. The addition (removal) of attractant causes a transient decrease (increase) in system activity, and thus of tumbling frequency. We observe a strong asymmetry in

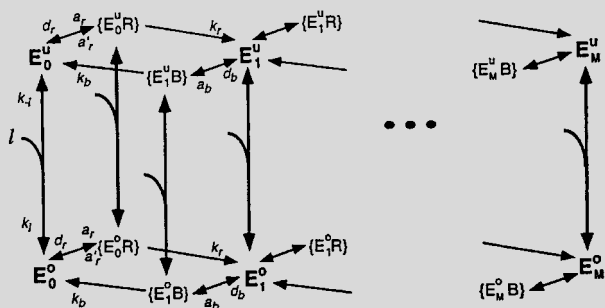
the response to the addition compared with the removal of ligand. This asymmetry has been observed experimentally¹⁴. The chemotactic response of the system has been measured by the average drift velocity in the presence of a linear gradient of attractant (Fig. 2, inset). The system is very sensitive: an average change in the receptor occupancy of $\sim 1\%$ per second is enough to induce a drift velocity of ~ 1 micron per second.

Figure 3a illustrates the most striking result of the model: we have found that the system shows almost perfect adaptation for a wide range of values of the network's biochemical parameters. Typically, one can change simultaneously each of the rate constants several-fold and still obtain, on average, only a few per cent deviation from perfect adaptation. For instance, over 80 per cent of model systems, obtained from a perfectly adaptive one by randomly changing all of its biochemical parameters by a factor of two, still show $<15\%$

Box 1 Two-state model of the bacterial chemotactic network

The main component of the two-state model^{2,19} is the receptor complex, MCP + CheA + CheW (Fig. 1a), considered here as a single entity, E. The complex is assumed to have two functional states—active and inactive. A receptor complex in the active state shows a kinase activity of CheA; by phosphorylating the response regulators, CheY, it sends a tumbling signal to the motors. The output of the network is thus the average number of receptors in the active state, the system activity A . It is assumed that this quantity determines the tumbling frequency of the bacteria. The transformation between A and the tumbling frequency depends on the kinetics of CheY phosphorylation and dephosphorylation, as well as on the interaction of CheY with the motors, which are not considered explicitly in the present model.

The receptor complexes are assumed to exist in different forms. Consider a complex methylated on m sites ($m = 1, \dots, M$). Such a complex can either be occupied or unoccupied by the ligand. We denote the concentration of these complexes by E_m^0 and E_m^u , respectively. Each form of the receptor complex can be in the active state with a probability depending on both its methylation level and its ligand occupancy. We assume that an occupied receptor complex has the probability α_m^0 of being in the active state; for an unoccupied receptor, this probability is α_m^u . If l is the ligand concentration, $B(R)$ the concentration of CheB (CheR), and $\{E_m^u B\}$ the concentration of the E_m^u CheB complex and so on, the model reactions can then be illustrated schematically (see figure).



The differential equations describing our model can be written in a standard way from the figure. For instance, the kinetic equation for E_m^u is

$$\begin{aligned} \frac{dE_m^u}{dt} = & -k_f l E_m^u + k_b E_m^0 + \\ & (1 - \delta_{m,0}) [-a_b \alpha_m E_m^u B + d_b \{E_m^u B\} + k_r \{E_{m-1}^u R\}] + \\ & (1 - \delta_{m,M}) [-a_r \alpha_m E_m^u R + a'_r (1 - \alpha_m) E_m^u R + d_b \{E_m^u R\} + k_b \{E_{m+1}^u B\}] \\ & (m = 1, \dots, M) \end{aligned}$$

The presence of α_m in the equation is due to the fact that CheB demethylates only the active receptors; we have also included two different association rate constants of CheR to the active (a_r) and the inactive (a'_r) receptors (see below). δ_k is the Kronecker's delta ($\delta_k = 1$, when $j = k$, and is zero otherwise). Similar equations can be written to describe kinetics of $\{E_m^u B\}$, $\{E_m^u R\}$, $\{E_m^0 B\}$ and $\{E_m^0 R\}$, with additional parameters α_m^0 defining the probabilities of E_m^0 to be in the active state. For fixed α_m and α_m^0 the biochemical parameters of this system include nine different rate constants ($k_f, k_b, a_r, a'_r, d_r, k_r, a_b, d_b, k_b$) and three enzyme concentrations (total concentrations of CheR, CheB and receptor complexes).

The present model is by no means the only two-state model of the chemotactic network that exhibits robust adaptation and proper chemotactic response. Rather, it is one of the simplest variants that is consistent with the experimental data on the response and adaptation of wild-type *E. coli*. The main assumptions underlying this model are as follows.

- The input to the system is the ligand concentration; rapid binding (and unbinding) of the ligand to the receptor induces an immediate change in the activity of the complex. For simplicity, the binding affinity is assumed to be independent of the receptor's activity and its degree of methylation. This assumption can be relaxed without affecting the main conclusions of our model.
- The methylation and demethylation reaction occur on slower timescales. A central assumption is that CheB can only demethylate active receptors. In addition, the demethylation rate constants do not depend explicitly either on ligand occupancy or on the methylation of the receptor, so that all active receptors are demethylated at the same rate. In the variant of the model discussed here, the phosphorylation of CheB is not considered explicitly; in molecular terms, we assume that the phosphorylated form of CheB, CheB_p, does not move freely in the cell. Rather, a CheB_p molecule can only demethylate the same receptor that has phosphorylated it. We note, however, that this assumption can also be readily relaxed (N.B. *et al.*, manuscript in preparation). Robust adaptation is maintained as long as both CheB and CheB_p demethylate only the active receptors.
- The methylating enzyme, CheR, acts both on active and inactive receptors. Here we assume that the association rate constant for this reaction depends only on the activity of the receptor (a'_r for inactive, a_r for active), whereas the dissociation rate constant d_r and the catalytic rate constant k_r are the same for all forms of the receptor. This assumption can again be relaxed in various ways. In particular, the accumulated biochemical data indicate that CheR works at saturation and operates at its maximal velocity. In this case, the conclusions of our model are not altered, even if d_r, a'_r and a_r depend on the ligand occupancy and on the methylation level²¹. □

deviation from perfect adaptation (Fig. 3a, lower panel). When varied separately, most of the rate constants may be changed by several orders of magnitude without inducing a significant deviation from perfect adaptation.

In our model we have assumed Michaelis–Menten kinetics for simplicity. However, we have found that cooperative effects in the enzymatic reactions can be added without destroying the robustness of adaptation. Similarly, robust adaptation is obtained for systems with different numbers of methylation sites. Multiple methylation sites are thus not required for robust adaptation, but possibly are for allowing strong initial responses for a wide range of attractant and repellent stimuli (N.B. *et al.*, manuscript in preparation).

The adaptation itself, as measured by its precision (Fig. 3a), is thus a robust property of the chemotactic network. This does not mean, however, that all the properties are equally insensitive to variations in the network parameters. For instance, Fig. 3b shows that the adaptation time, τ , which characterizes the dynamics of relaxation to the steady-state activity, displays substantial variations in the altered systems. Robustness is thus a characteristic of specific network properties and not of the network as a whole: whereas some properties are robust, others can show sensitivity to changes in the network parameters.

Plots similar to the ones depicted in Fig. 3 can be obtained in quantitative experiments. A large collection of chemotactic mutants can be analysed for variations in the biochemical rate constants of the chemotactic network components. Alternatively, the rate constants of the enzymes could be systematically modified or their expression varied. At the same time, their various physiological characteristics can be measured, such as steady-state tumbling frequency, precision of adaptation, adaption time, and so on. In this way, the predictions of the model can be quantitatively checked.

What features of the chemotactic network make the adaptation property so robust? We propose here a general and simple mechanism for robust adaptation. Let us introduce this mechanism for one of the simplest networks (Fig. 1b), which can be viewed either as an ‘adaptation module’, or, as a simplifying reduction of a more complex adaptive network, such as the one presented for bacterial chemotaxis. Consider an enzyme, E, which is sensitive to an external signal l , such as a ligand. Each enzyme molecule is at equilibrium between two functional states: an active state, in which it catalyses a reaction, and an inactive state, in which it does not. The signal level l affects the equilibrium between two functional states of the enzyme: we suppose that a change in l causes a rapid response of the system by shifting this equilibrium. Thus, l is the input of this signal transduction system and the concentration of active enzymes (that is, the system activity, A) can be considered as its output. The enzyme E can be reversibly modified, for example by addition of methyl or phosphate groups. The modification of E affects the probabilities of the active and inactive states, and hence can compensate for the effect of the ligand. In general, then, $A(l) = \alpha(l)E + \alpha_m(l)E_m$, where E_m and E are the concentrations of the modified and unmodified enzyme, respectively, and $\alpha_m(l)$ and $\alpha(l)$ are the probabilities that the modified and unmodified enzyme is active. After an initial rapid response of the system to a change in the input level, l , slower changes in the system activity proceed according to the kinetics of enzyme modification.

The system is adaptive when its steady-state activity, A^{st} , is independent of l . A mechanism for adaptation can be readily obtained by assuming a fine-tuned dependence of the biochemical parameters on the signal level, l . This kind of mechanism has been proposed for an equivalent receptor system^{17,18}. A mechanism for robust adaptation, on the other hand, can be obtained when the rates of the modification and the reverse-modification reactions depend solely on the system activity, A , and not explicitly on the concentrations E_m and E . This system can be viewed as a feed-back system, in which the output A determines the rates of modification

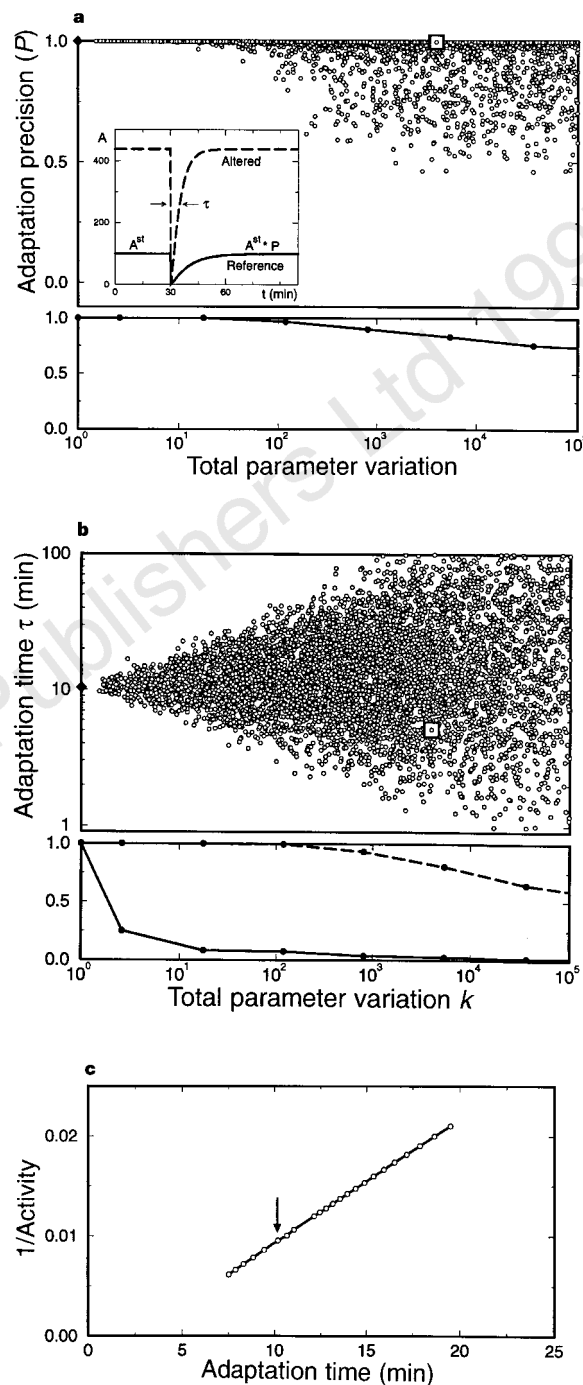


Figure 3 Robustness of adaptation. **a**, The precision of adaptation, P , and **b**, adaptation time, τ , to a step-like addition of saturating amount of attractant are plotted as a function of the total parameter variation k , for an ensemble of model systems (see Methods). The time evolution of the system activity A is depicted in the inset for the reference system (solid curve) and for an altered model system, obtained by randomly increasing or decreasing by a factor of two all biochemical parameters of the reference system (dashed curve). Each point in the top graphs in **a** and **b** corresponds to a different altered system, out of the total number of 6,157. The reference system is denoted by a black diamond; the particular altered system from the inset is denoted by an open square. Bottom graphs: **a**, the probability that P is larger than 0.95; **b**, the probability that τ deviates from the adaptation time of the reference system (~ 10 min) by less than 5% (solid curve) and by a factor 5 (dashed curve). **c**, ‘Individuality’ in the chemotaxis model. The inverse steady-state activity A^{-1} is plotted as a function of the adaptation time, τ . Each point represents an altered system, obtained from the reference system (arrow) by varying the concentration of CheR (between 100 to 300 molecules per cell).

reactions, which in turn determine the slow changes in A . With such activity-dependent kinetics, the value of the steady-state activity, A^{st} , is independent of the ligand level, therefore the system is adaptive. Activity-dependent kinetics can be achieved in a variety of ways. As a simple example, consider a system for which only the modified enzyme can be active ($\alpha = 0$); the enzyme R , which catalyses the modification reaction $E \rightarrow E_m$, works at saturation, and the enzyme B , which catalyses the reverse-modification reaction $E_m \rightarrow E$, can only bind to active enzymes. In this case, the modification rate is constant at all times, whereas the reverse modification rate is a simple function of the activity

$$\frac{dE_m}{dt} = V_{\max}^R - V_{\max}^B \frac{A}{K_b + A} \quad (1)$$

where $V_{\max}^B$ and $V_{\max}^R$ are the maximal velocities of the modification and the reverse-modification reactions, respectively, and K_b is the Michaelis constant for the reverse modification reaction; we have assumed $V_{\max}^R < V_{\max}^B$. For simplicity, we have assumed that the enzymes follow Michaelis–Menten (quasi-steady-state) kinetics. The functioning of the feedback can now be analysed: the system activity is continuously compared to a reference steady-state value

$$A^{st} = K_b \frac{V_{\max}^R}{V_{\max}^B - V_{\max}^R}.$$

For $A < A^{st}$, the amount of modification increases, leading to an increase in A ; for $A > A^{st}$, the modification decreases, leading to a decrease in A . In this way, the system always returns to its steady-state value of activity, exhibiting adaptation. Moreover, with these activity-dependent kinetics, the adaptation properties is insensitive to the values of system parameters (such as enzyme concentrations), so adaptation is robust.

Note, however, that the steady-state activity itself, which is not a robust property of the network, depends on the enzyme concentrations. Thus, the mechanism presented here still provides a way to control the system activity on long timescales, for example by changing the expression level of the modifying enzymes while preserving adaptation itself on shorter timescales.

A quantitative analysis demonstrates that, on methylation time-scales, the kinetics of the two-state model of chemotaxis can, for a wide range of parameters, be mathematically ‘reduced’ to the simple activity-dependent kinetics shown in equation (1) (N.B. *et al.*, manuscript in preparation). Robust adaptation thus follows naturally as consequence of the simple mechanism described above. The deviations from perfect adaptation (Fig. 3) are in fact connected to departures from the assumptions underlying this mechanism (such as $V_{\max}^R < V_{\max}^B$). This simple mechanism suggests that the various detailed assumptions about the system’s biochemistry can be easily altered, provided that the activity-dependent kinetics of receptor modifications is preserved. All variants of the model obtained in this way still exhibit robust adaptation (N.B. *et al.*, manuscript in preparation).

Two main observations argue in favour of a robust, rather than a fine-tuned, adaptation mechanism for chemotaxis. First, the adaptation property is observed in a large variety of chemotactic bacterial populations. It is easier to imagine how a robust mechanism allows bacteria to tolerate genetic polymorphism, which may change the network’s biochemical parameters. In addition, in genetically identical bacteria some features of the chemotactic response, such as the values of adaptation time and of steady-state tumbling frequency, vary significantly from one bacterium to another, while the adaptation property itself is preserved²⁰. This ‘individuality’ can be readily explained in the framework of the present model. The concentrations of some cellular proteins, for example the methylating enzyme CheR, are very low², and thus may

be subject to considerable stochastic variations. In consequence, both adaptation time and steady-state tumbling frequency, which are not robust properties of the network, should vary significantly. Moreover, the present model predicts that both these quantities should show a strong correlation in their variation (Fig. 3c), which has been observed experimentally²⁰.

How general are the results presented here? In addition to explaining response and adaptation in chemotaxis, the present model accounts, in a unifying way, for other taxis behaviour of bacteria mediated by the same network. Indeed, as the network’s dynamics is solely determined by the system activity, the system will respond and adapt to any environmental change that affects this activity. Mechanisms of robust adaptation similar to the one introduced above could apply to a wider class of signal transduction networks. Robustness may be a common feature of many key cellular properties and could be crucial for the reliable performance of many biochemical networks. Robust properties of a network will be preserved even if its components are modified through random mutations, or are produced in modified quantities. Systems whose key properties are robust could have an important advantage in having a larger parameter space in which to evolve and to adjust to environmental changes.

The degree of robustness in many biochemical networks can be quantitatively investigated. This can be achieved by characterizing a behavioural, a physical or biochemical property while varying systematically the expression level and the rate constants of the network’s components.

The complexity of biological systems introduce several conceptual and practical difficulties, however. Among the most important is the difficulty of isolating smaller subsystems that could be analysed separately. For instance, in the present analysis, we have neglected the existence of different types of receptors and any crosstalk between them. We have also disregarded the interactions between the chemotaxis network and other components of the cell. In addition, the complexity and stochastic variability of biological networks may preclude their complete molecular description. Rate constants and concentrations of many enzymes can only be measured outside their natural cellular environment and many other network parameters remain unknown. Robustness may provide a way out of both these quandaries: robust properties do not depend on the exact values of the network’s biochemical parameters and should be relatively insensitive to the influence of the other subsystems. It should then be possible to extract some of the principles underlying cell function without a full knowledge of the molecular detail. □

Methods

Numerical integration of the kinetics equations defining the two-state model (see Box 1) was used to investigate its properties. Computer programs in C++ language were executed on an SGI (R4000) workstation using a standard routine (*lsode* from LLNL). Typical CPU time for finding a numerical solution of a model system is of the order of 1 min. A particular model system was obtained by assigning values to the rate constants and the total enzyme concentrations. Most of our results were obtained for a reference system defined by the following biochemical parameters: the equilibrium binding constant of ligand to receptor is $1 \mu\text{M}$ and the time constant for the reaction is 1 ms ($k_1 = 1 \text{ ms}^{-1} \mu\text{M}^{-1}$, $k_{-1} = 1 \text{ ms}^{-1}$). CheR methylates both active and inactive receptors at the same rate, with a Michaelis constant of $1.25 \mu\text{M}$, and a time constant of 10 s ($a_r = a'_r = 80 \text{ s}^{-1} \mu\text{M}^{-1}$, $d_r = 100 \text{ s}^{-1}$, $k_r = 0.1 \text{ s}^{-1}$). CheB (CheB_p) demethylates only active receptors with a Michaelis constant of $1.25 \mu\text{M}$ and a time constant of 10 s ($a_b = 800 \text{ s}^{-1} \mu\text{M}^{-1}$, $d_b = 1,000 \text{ s}^{-1}$, $k_b = 0.1 \text{ s}^{-1}$). The number of enzyme molecules per cell are: 10,000 receptor complexes, 2,000 CheB and 200 CheR (cell volume of $1.4 \times 10^{-15} \text{ l}$). The probabilities that a receptor with $m = 1, \dots, 4$ methylated sites is in its active state are: $\alpha_1 = 0.1$, $\alpha_2 = 0.5$, $\alpha_3 = 0.75$, $\alpha_4 = 1$ if it is unoccupied, and $\alpha_1^0 = 0$, $\alpha_2^0 = 0.1$, $\alpha_3^0 = 0.5$, $\alpha_4^0 = 1$ if it is occupied.

Response and adaptation. In a typical assay, a model system was subject to a step-like change in attractant concentration. A system in steady-state, characterized by the system activity A^{st} , was perturbed by an addition or removal of attractant. As a result, the system activity changed abruptly and then relaxed, with the characteristic adaptation time, τ , to a new steady-state value $A^{st} \cdot p$. Here p measures the precision of adaptation; perfect adaptation corresponds to $p = 1$ (see inset in Fig. 3a).

Robustness of adaptation. The sensitivity of adaptation precision and adaptation time to variations in the biochemical constants defining a model system was investigated. An ensemble of altered systems was obtained from the reference system by random modifications of its reaction rate constants and enzymatic concentrations, k_n^0 . Each alteration of the reference system was characterized by the total parameter variation, k , which is defined as: $\log(k) = \sum_{n=1}^N |\log(k_n/k_n^0)|$, where k_n are the biochemical parameters of the altered system. The altered system was subject to a step-like addition of saturating concentrations of attractant (1 mM), and both the precision of adaptation, p , and the adaptation time, τ , were measured. The assay was repeated for various reference model systems, with different values of biochemical parameters and of α_m , and different variants of the model. The robustness of adaptation (Fig. 3) is independent of these choices.

Chemotactic drift velocity. The behaviour of a model system in the presence of a linear gradient of attractant, ∇I , was simulated. The movement of the system was assumed to be composed of a series of smooth runs at a constant velocity of $20 \mu\text{m s}^{-1}$, interrupted by tumbling events. The tumbling frequency was taken to be a sigmoidal function of the system activity (Hill coefficient, $q = 2$). Different values of q lead to the same qualitative picture; the sensitivity increases with q . The trajectories were also subject to a rotation diffusion, with $D = 0.125 \text{ rad}^2 \text{ s}^{-1}$ (ref. 9). Attractant concentration was increasing along the x direction, (with $l = 1 \mu\text{M}$ at $x = 0$). The chemotactic drift velocity was estimated by measuring the average x position of a hundred identical simulated systems as a function of time.

Received 31 December 1996; accepted 17 April 1997.

- Bray, D. Protein molecules as computational elements in living cells. *Nature* **376**, 307–312 (1995).
- Stock, J. B. & Surette, M. in *E. coli and S. typhimurium: Cellular and Molecular Biology* (ed. Neidhardt, F. C.) 1103–1129 (American Society of Microbiology, Washington DC, 1996).
- Parkinson, J. S. Signal transduction schemes of bacteria. *Cell* **73**, 857–871 (1993).
- Hazelbauer, G. L., Berg, H. C. & Matsumura, P. M. Bacterial motility and signal transduction. *Cell* **73**, 15–22 (1993).
- Bourret, R. B., Borkovich, K. A. & Simon, M. I. Signal transduction pathways involving protein phosphorylation in prokaryotes. *Annu. Rev. Biochem.* **60**, 401–441 (1991).
- Adler, J. Chemotaxis in bacteria. *Annu. Rev. Biochem.* **44**, 341–356 (1975).
- Bray, D., Bourret, R. B. & Simon, M. I. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Mol. Biol. Cell* **4**, 469–482 (1993).
- Adler, J. Chemotaxis in bacteria. *Science* **153**, 708–716 (1996).
- Berg, H. C. & Brown, D. A. Chemotaxis in *E. coli* analysed by three-dimensional tracking. *Nature* **239**, 500–504 (1972).
- Macnab, R. M. & Koshland, D. E. The gradient-sensing mechanism in bacterial chemotaxis. *Proc. Natl Acad. Sci. USA* **69**, 2509–2512 (1972).
- Block, S. M., Segall, J. E. & Berg, H. C. Impulse responses in bacterial chemotaxis. *Cell* **31**, 215–226 (1982).
- Koshland, D. E. A response regulator model in a simple sensory system. *Science* **196**, 1055 (1977).
- Berg, H. C. & Tedesco, P. M. Transient response to chemotactic stimuli in *E. coli*. *Proc. Natl Acad. Sci. USA* **72**, 3235–3239 (1975).
- Springer, M. S., Goy, M. F. & Adler, J. Protein methylation in behavioural control mechanism and in signal transduction. *Nature* **280**, 279–284 (1979).
- Koshland, D. E., Goldbeter, A. & Stock, J. B. Amplification and adaptation in regulatory and sensory systems. *Science* **217**, 220–225 (1982).
- Khan, S., Spudich, J. L., McCray, J. A. & Tentham, D. R. Chemotactic signal integration in bacteria. *Proc. Natl Acad. Sci. USA* **92**, 9757–9761 (1995).
- Segel, I. A., Goldbeter, A., Devreotes, P. N. & Knox, B. E. A mechanism for exact sensory adaptation based on receptor modification. *J. Theor. Biol.* **120**, 151–179 (1986).
- Hauri, D. C. & Ross, J. A model of excitation and adaptation in bacterial chemotaxis. *Biophys. J.* **68**, 708–722 (1995).
- Asakura, S. & Honda, H. Two-state model for bacterial chemoreceptor proteins. *J. Mol. Biol.* **176**, 349–367 (1984).
- Spudich, J. L. & Koshland, D. E. Non-genetic individuality: chance in the single cell. *Nature* **262**, 467–471 (1976).
- Kleene, S. J., Hobson, A. C. & Adler, J. Attractants and repellents influence methylation and demethylation of methyl-accepting proteins in an extract of *E. coli*. *Proc. Natl Acad. Sci. USA* **76**, 6309–6313 (1979).

Acknowledgements. We thank J. Stock, M. Surette, A. C. Maggs, U. Alon, L. Hartwell, M. Kirschner, A. Levine, A. Libchaber, A. Murray and T. Surrey for discussion; A. C. Maggs for help with numerical issues; and J. Stock, M. Surette and H. Berg for introducing us to bacterial chemotaxis and pointing out many useful references. This work has been partially supported by grants from the NIH and the NSF. N.B. is a Rothschild Fellow and a Dicke Fellow at Princeton University.

Correspondence and requests for materials should be addressed to S.L. (e-mail: leibler@princeton.edu).

A family of cytokine-inducible inhibitors of signalling

Robyn Starr*, Tracy A. Willson*, Elizabeth M. Viney*, Leeclia J. L. Murray*, John R. Rayner†, Brendan J. Jenkins†, Thomas J. Gonda†, Warren S. Alexander*, Donald Metcalf*, Nicos A. Nicola* & Douglas J. Hilton*

* The Walter and Eliza Hall Institute for Medical Research and The Cooperative Research Center for Cellular Growth Factors, Parkville, Victoria, Australia 3052

† The Hanson Centre for Cancer Research, IMVS, Adelaide, Southern Australia, Australia 5000

Cytokines are secreted proteins that regulate important cellular responses such as proliferation and differentiation¹. Key events in cytokine signal transduction are well defined: cytokines induce receptor aggregation, leading to activation of members of the JAK family of cytoplasmic tyrosine kinases. In turn, members of the STAT family of transcription factors are phosphorylated, dimerize and increase the transcription of genes with STAT recognition sites in their promoters^{1–4}. Less is known of how cytokine signal transduction is switched off. We have cloned a complementary DNA encoding a protein SOCS-1, containing an SH2-domain, by its ability to inhibit the macrophage differentiation of M1 cells in response to interleukin-6. Expression of SOCS-1 inhibited both interleukin-6-induced receptor phosphorylation and STAT activation. We have also cloned two relatives of SOCS-1, named SOCS-2 and SOCS-3, which together with the previously described CIS (ref. 5) form a new family of proteins. Transcription of all four SOCS genes is increased rapidly in response to interleukin-6, *in vitro* and *in vivo*, suggesting they may act in a classic negative feedback loop to regulate cytokine signal transduction.

To identify cDNAs encoding proteins capable of suppressing cytokine signal transduction, we used an expression cloning approach. The strategy used the murine monocytic leukaemic M1 cell line that differentiates into mature macrophages and ceases proliferation in response to various cytokines, including interleukin-6 (IL-6), and in response to the steroid, dexamethasone^{6,7}. Parental M1 cells were infected with the RUFneo retrovirus, into which a library of cDNAs from the factor-dependent haemopoietic cell line FDC-P1 had been inserted⁸. Retrovirally infected M1 cells that were unresponsive to IL-6 were selected in semi-solid agar culture by their ability to generate compact colonies in the presence of IL-6 and geneticin. One stable IL-6-unresponsive clone, 4A2, was obtained after examining 10^4 infected cells (Fig. 1). A 1.4 kilobase pair (kbp) cDNA insert, which we have named suppressor of cytokine signalling-1, or SOCS-1, was recovered by polymerase chain reaction (PCR) from the retrovirus that had integrated into genomic DNA of 4A2 cells. The SOCS-1 PCR product was used to

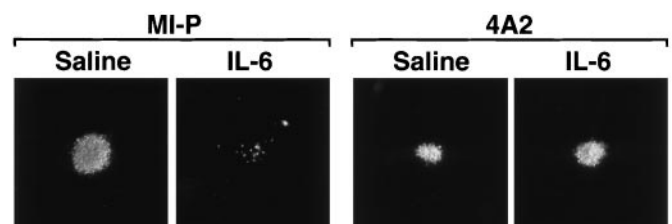
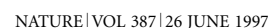


Figure 1 Phenotype of IL-6 unresponsive M1 cell clone, 4A2. Colonies of parental M1 cells (left panel) and clone 4A2 (right panel) cultured in semi-solid agar for 7 days in saline or 100 ng ml^{-1} IL-6.

To establish that the phenotype of the 4A2 cell line was directly related to expression of SOCS-1, and not to unrelated genetic changes that may have occurred in this cell clone, a cDNA encoding an epitope-tagged version of SOCS-1 was cloned into an expression vector under the control of the EF1 α promoter. The SOCS-1 expression vector was cotransfected with a plasmid conferring resistance to puromycin into parental M1 cells and, in order to assess the effect of SOCS-1 expression on signalling by a broader



range of cytokines, into M1 cells expressing the receptor for thrombopoietin, c-mpl (M1.mpl). Multiple independent clones of M1 cells expressing SOCS-1, as detected by western blot, displayed a cytokine-unresponsive phenotype that was indistinguishable from 4A2 (data not shown). Further, if transfectants were not maintained in puromycin, expression of SOCS-1 was lost over time and cells regained their cytokine responsiveness. In the absence of cytokine, colonies derived from 4A2 and other SOCS-1-expressing clones characteristically grew to a smaller size than colonies formed by control M1 cells (Fig. 1).

The effect of constitutive SOCS-1 expression on the response of M1 cells to a range of cytokines was further investigated using the 4A2 cell line and a clone of M1.mpl cells expressing SOCS-1 (M1.mpl.SOCS-1). Unlike parental M1 cells and M1.mpl cells, the two cell lines expressing SOCS-1 continued to proliferate and failed to form differentiated colonies in response to either IL-6, leukaemia inhibitory factor (LIF), oncostatin M (OSM), interferon- γ (IFN- γ) or, in the case of the M1.mpl.SOCS-1 cell line, thrombopoietin (Fig. 3). However, both cell lines were able to respond to dexamethasone, suggesting that SOCS-1 specifically affected cytokine signal transduction rather than the cell's intrinsic ability to differentiate. Consistent with these data, 4A2 and M1.mpl.SOCS-1 cells showed no evidence of morphological differentiation in response to IL-6 or other cytokines, whereas parental M1 cells and M1.mpl cells became large and vacuolated with a macrophage morphology in response to IL-6 (data not shown).

Because CIS was cloned as a cytokine-inducible immediate early gene, we examined whether SOCS-1, SOCS-2 and SOCS-3 were similarly regulated. The basal pattern of expression of the four SOCS genes was examined by northern blot analysis of messenger RNA from a variety of tissues from male and female C57BL/6 mice (Fig. 4a). Constitutive expression of SOCS-1 was observed in the thymus and to a lesser extent in the spleen and the lung. SOCS-2 expression was restricted primarily to the testis and in some animals the liver and lung; for SOCS-3 a low level of expression was observed

in the lung, spleen and thymus, whereas CIS expression was more widespread, including the testis, heart, lung, kidney and, in some animals, the liver.

We sought to determine whether expression of the four SOCS genes was regulated by IL-6. Northern blots of mRNA prepared from the livers of untreated and IL-6-injected mice, or from unstimulated and IL-6-stimulated M1 cells, were hybridized with labelled fragments of SOCS-1, SOCS-2, SOCS-3 and CIS cDNAs (Fig. 4b). Expression of all four SOCS genes was increased in the liver following IL-6 injection; however, the kinetics of induction appeared to differ. Expression of SOCS-1 and SOCS-3 was transient in the liver, with mRNA detectable 20 min after IL-6 injection and declining to basal levels within 4 h for SOCS-1 and 8 h for SOCS-3. Induction of SOCS-2 and CIS mRNA in the liver followed similar initial kinetics to that of SOCS-1, but was maintained at an elevated level for at least 24 h. A similar induction of SOCS gene mRNA was observed in other organs, notably the lung and the spleen (data not shown). In contrast, in M1 cells although SOCS-1 and CIS mRNA were induced by IL-6, no induction of either SOCS-2 or SOCS-3 expression was detected. This result highlights cell-type-specific differences in the expression of the genes of SOCS family members in response to the same cytokine.

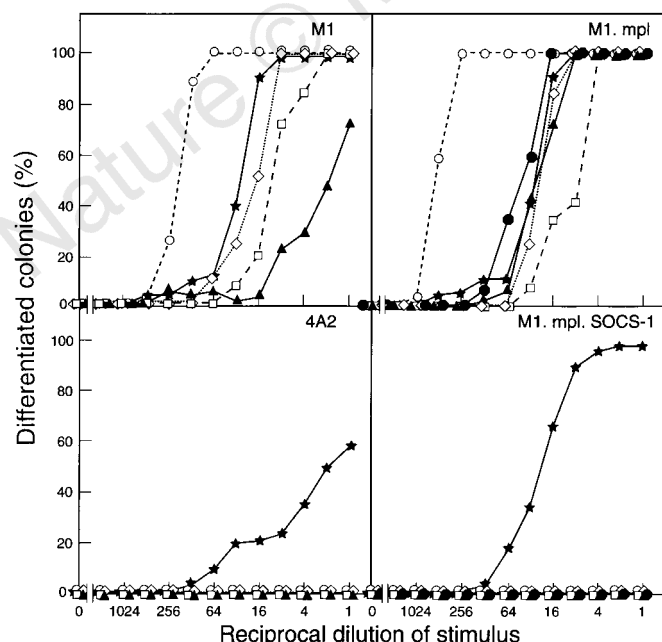


Figure 3 Expression of SOCS-1 suppresses IL-6-induced macrophage differentiation of M1 cells. Semi-solid agar cultures of parental M1 cells (M1 and M1.mpl) and M1 cells expressing SOCS-1 (4A2 and M1.mpl.SOCS-1), showing the percentage of colonies that differentiated in response to a titration of $1 \mu\text{g ml}^{-1}$ IL-6 (empty circles), 100 ng ml^{-1} LIF (empty triangles), $1 \mu\text{g ml}^{-1}$ OSM (squares), 100 ng ml^{-1} IFN- γ (filled triangles), 500 ng ml^{-1} thrombopoietin (TPO) (filled circles) or $3 \times 10^{-6} \text{ M}$ dexamethasone (stars).

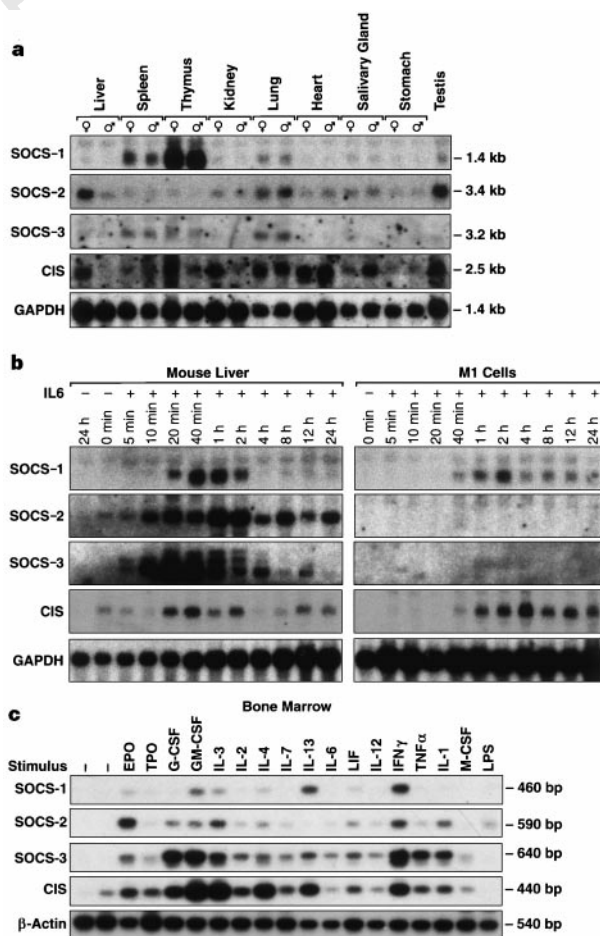


Figure 4 Expression of mRNA for SOCS family members *in vitro* and *in vivo*. **a**, Northern analysis of mRNA from a range of mouse organs showing constitutive expression of SOCS family members in a limited number of tissues. **b**, Northern analysis of mRNA from liver and M1 cells showing induction of expression of SOCS family members following exposure to IL-6. **c**, Reverse transcriptase PCR analysis of mRNA from bone marrow showing induction of expression of SOCS family members by a range of cytokines.

In order to examine the spectrum of cytokines that was capable of inducing transcription of the various members of the SOCS gene family, bone marrow cells were stimulated for an hour with a range of cytokines, after which mRNA was extracted and cDNA was synthesized. PCR was then used to assess the expression of SOCS-1, SOCS-2, SOCS-3 and CIS (Fig. 4c). In the absence of stimulation, little or no expression of any of the SOCS genes was detectable in bone marrow by PCR. Stimulation of bone marrow cells with a broad array of cytokines appeared capable of upregulating mRNA for one or more members of the SOCS family. IFN- γ , for example, induced expression of all four SOCS genes, whereas erythropoietin, granulocyte colony-stimulating factor, granulocyte-macrophage colony-stimulating factor and interleukin-3 induced expression of SOCS-2, SOCS-3 and CIS. Interestingly, tumour necrosis factor- α , macrophage colony-stimulating factor and interleukin-1, which act through receptors that do not fall into the type I cytokine receptor class also appeared capable of inducing expression of SOCS-3 and CIS, suggesting that the SOCS proteins may play a broader role in regulating signal transduction.

As constitutive expression of SOCS-1 inhibited the response of M1 cells to a range of cytokines, we examined whether phosphorylation of the cell-surface receptor component gp130 and the

transcription factor Stat3, which are thought to play a central role in IL-6 signal transduction¹⁻⁴, were affected. We compared these events in the parental M1 and M1.mpl cell lines and their SOCS-1-expressing counterparts. As expected, gp130 was phosphorylated rapidly in response to IL-6 in both parental lines, however, this was reduced in the cell lines expressing SOCS-1 (Fig. 5a). Likewise, Stat3 phosphorylation was also reduced in response to IL-6 in those cell lines expressing SOCS-1 (Fig. 5a). Consistent with a reduction in Stat3 phosphorylation, activation of specific STAT/DNA-binding complexes, as determined by electrophoretic mobility shift assay, was also reduced. Notably, there was a failure to form SIF-A (containing Stat3) and SIF-B (Stat1/Stat3 heterodimer), the major STAT complexes induced in M1 cells stimulated with IL-6 (Fig. 5b). Similarly, constitutive expression of SOCS-1 also inhibited IFN- γ -stimulated formation of SIF-C (Stat1 homodimer; Fig. 5b).

The effects of SOCS-1 appeared to be specific for the JAK/STAT pathway because no detectable diminution of the overall level of tyrosine-phosphorylated proteins was observed either in the absence or presence of IL-6 in M1 cells expressing SOCS-1 (data not shown). Similarly in M1 cells, which express flt3/flk-2 (ref. 12), flt3/flk-2 ligand-induced tyrosine phosphorylation of its receptor and the cytoplasmic protein Shc were unaffected by expression of SOCS-1 (data not shown). Taken together, these experiments are consistent with the proposal that SOCS-1 specifically inhibits signal transduction upstream of receptor and STAT phosphorylation. SOCS-1 may therefore act as a specific inhibitor of JAK kinases, and because these enzymes have been implicated in the transmission of both differentiative and proliferative signals¹³, it is possible that SOCS-1 may also inhibit cytokine-induced cell proliferation.

The ability of SOCS-1 to inhibit signal transduction and ultimately the biological response to cytokines suggests that, like the SH2-containing phosphatase SHP-1 (refs 14, 15), the SOCS proteins play a central role in controlling the intensity and/or duration of a cell's response to a diverse range of extracellular stimuli by suppressing the signal transduction process. The evidence provided here and in accompanying papers suggests that the SOCS family acts in a classic negative feedback loop for cytokine signal transduction. Like other genes such as OSM, expression of genes encoding the SOCS proteins is induced by cytokines through the activation of STATs^{16,17}. Once expressed, it is proposed that the SOCS proteins inhibit the activity of JAKs and so reduce the phosphorylation of receptors and STATs, thereby suppressing signal transduction and any ensuing biological response. Importantly, inhibition of STAT activation will, over time, lead to a reduction in SOCS gene expression, allowing cells to regain responsiveness to cytokines. Assessment of the physiological role of SOCS-1, SOCS-2, SOCS-3 and CIS will be clarified further by generating mice that lack one or more members of this family.

Note added in proof: Elsewhere in this issue, JAK inhibitors similar to SOCS are reported under the names JAB¹¹ and SSI-1 (ref. 27). □

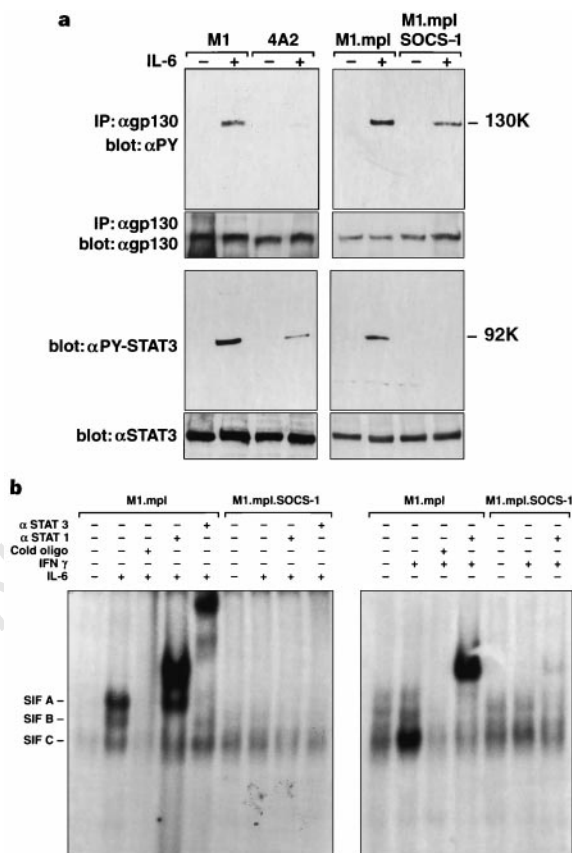


Figure 5 SOCS-1 suppresses the phosphorylation and activation of gp130 and Stat-3. **a**, Western blots of extracts from parental M1 cells (M1 and M1.mpl) and M1 cells expressing SOCS-1 (4A2 and M1.mpl.SOCS-1) stimulated with (+) or without (-) 100 ng ml⁻¹ IL-6. Top: extracts immunoprecipitated with anti-gp130 (α gp130) and immunoblotted with anti-phosphotyrosine (α PY). The blot was reprobed with α gp130 to demonstrate equal loading of protein. Bottom: cell lysates (10 μ g) immunoblotted either with an antibody specific for tyrosine phosphorylated-STAT3 (α PY-Stat3), or for Stat3 (α Stat3) to demonstrate equal loading of protein. The molecular masses of the bands are shown on the right. **b**, EMSA of M1.mpl and M1.mpl.SOCS-1 cells stimulated with (+) and without (-) 100 ng ml⁻¹ IL-6 or 100 ng ml⁻¹ IFN- γ . The DNA-binding complexes SIFA, B, and C are indicated at the left.

Methods

Generation of retroviral library and infection of M1 cells. A cDNA expression library of 10⁴ clones was constructed from the factor-dependent haemopoietic cell line FDC-P1 and used to infect M1 cells, essentially as described⁸. Infected IL-6 unresponsive M1 cells were selected in agar cultures containing 400 μ g ml⁻¹ geneticin and 100 ng ml⁻¹ IL-6.

Database searches and cloning of cDNAs. cDNA inserts were recovered from integrated retroviruses by PCR, as described⁸. Independent cDNA clones encoding mouse SOCS-1 were isolated from a murine thymus cDNA library essentially as described¹⁸. The nucleotide and predicted amino-acid sequence of mouse SOCS-1 cDNA was compared to databases using the BLASTN and TFASTA algorithms¹⁹⁻²¹. Oligonucleotides were designed from the ESTs encoding human SOCS-1 and mouse SOCS-2 and SOCS-3 and used to probe commercially available mouse thymus and spleen cDNA libraries essentially as described previously¹⁸. Sequencing was performed using an ABI automatic sequencer according to the manufacturer's instructions.

Northern analysis and reverse transcription. Purification of poly(A)⁺ mRNA, and northern blot hybridization were performed essentially as described²². To assess the induction of SOCS genes by IL-6, mice (C57BL/6) were injected intravenously with 5 µg IL-6 followed by collection of the liver at the indicated timepoints after injection. M1 cells were cultured in the presence of 20 ng ml⁻¹ IL-6 and collected at the indicated times. For RT-PCR analysis, bone marrow cells were collected as described²³ and stimulated for 1 h at 37 °C with 100 ng ml⁻¹ of a range of cytokines. RT-PCR was performed on total RNA as described²³. PCR products were resolved on an agarose gel and Southern blots were hybridized as described²³ with probes specific for each SOCS family member. Expression of β-actin was assessed to ensure uniformity of amplification.

Western blotting and electrophoretic mobility shift assays. M1 cells (10⁷) or their derivatives were stimulated for 4 min at 37 °C with either saline or 100 ng ml⁻¹ IL-6. Cells were lysed and 1 to 2 mg protein was immunoprecipitated with 4 µg anti-gp130 antibody (M20; Santa Cruz Biotechnology, Santa Cruz, CA) essentially as described²⁴. Western blots were performed using anti-tyrosine phosphorylated STAT3 or anti-STAT3 (New England Biolabs, Beverly, MA), or anti-gp130 (Santa Cruz Biotechnology) as described²⁵. Electrophoretic mobility shift assays were performed using the m67 oligonucleotide probe, as described²⁶.

Received 19 December 1996; accepted 21 March 1997.

- Nicola, N. A. *Guidebook to Cytokines and Their Receptors* (Oxford Univ. Press, Oxford, 1994).
- Ihle, J. N. Cytokine receptor signalling. *Nature* **377**, 591–594 (1995).
- Ihle, J. N., Witthuhn, B. A., Quelle, F. W., Yamamoto, K. & Silvennoinen, O. Signaling through the hematopoietic cytokine receptors. *Ann. Rev. Immunol.* **13**, 369–398 (1995).
- Darnell, J. Jr, Kerr, I. M. & Stark, G. R. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264**, 1415–1421 (1994).
- Yoshimura, A. *et al.* A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosine-phosphorylated interleukin 3 and erythropoietin receptors. *EMBO J.* **14**, 2816–2826 (1995).
- Metcalfe, D., Hilton, D. J. & Nicola, N. A. Clonal analysis of the actions of the murine leukemia inhibitory factor on leukemic and normal murine hematopoietic cells. *Leukemia* **2**, 216–221 (1988).
- Lotem, J., Shabo, Y. & Sachs, L. Clonal variation in susceptibility to differentiation by different protein. *Leukemia* **3**, 804–807 (1989).
- Rayner, J. R. & Gonda, T. J. A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. *Mol. Cell. Biol.* **14**, 880–887 (1994).
- Schlueter, G. *et al.* Sequence analysis of the conserved Protamine gene cluster shows that it contains a fourth expressed gene. *Mol. Reprod. Dev.* **43**, 1–6 (1996).
- Feng, J. *et al.* Activation of JAK2 catalytic activity requires phosphorylation of Y1007 in the kinase activation loop. *Mol. Cell. Biol.* (in the press).
- Endo, T. A. *et al.* A new protein containing an SH2 domain that inhibits JAK kinases. *Nature* **387**, 921–924 (1997).
- Rasko, J. E., Metcalfe, D., Rossner, M. T., Begley, C. G. & Nicola, N. A. The flt3/flk-2 ligand: receptor distribution and action on murine haemopoietic cell survival and proliferation. *Leukemia* **9**, 2058–2066 (1995).
- Smith, A., Metcalfe, D. & Nicola, N. A. Cytoplasmic domains of the common β-chain of the GM-CSF/IL-3/IL-5 receptors that are required for inducing differentiation or clonal suppression in myeloid leukaemic cell lines. *EMBO J.* **16**, 451–464 (1997).
- Ihle, J. N. *et al.* Protein tyrosine phosphorylation in the regulation of hematopoiesis by receptors of the cytokine-receptor superfamily. *Blood Cells* **20**, 65–80 (1994).
- Yi, T., Mui, A. L., Krystal, G. & Ihle, J. N. Hematopoietic cell phosphatase associates with the interleukin-3 (IL-3) receptor beta chain and down-regulates IL-3-induced tyrosine phosphorylation and mitogenesis. *Mol. Cell. Biol.* **13**, 7577–7586 (1993).
- Mui, A. L., Wakao, H., Kinoshita, T., Kitamura, T. & Miyajima, A. Suppression of interleukin-3-induced gene expression by a C-terminal truncated Stat5: role of Stat5 in proliferation. *EMBO J.* **15**, 2425–2433 (1996).
- Quelle, F. W. *et al.* Erythropoietin induces activation of Stat5 through association with specific tyrosines on the receptor that are not required for a mitogenic response. *Mol. Cell. Biol.* **16**, 1622–1631 (1996).
- Hilton, D. J. *et al.* Cloning of a murine IL-11 receptor alpha-chain; requirement for gp130 for high affinity binding and signal transduction. *EMBO J.* **13**, 4765–4775 (1994).
- Pearson, W. R. & Lipman, D. J. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448 (1988).
- Pearson, W. R. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* **183**, 63–98 (1990).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- Alexander, W. S., Metcalfe, D. & Dunn, A. R. Point mutations within a dimer interface homology domain of c-Mpl induce constitutive receptor activity and tumorigenicity. *EMBO J.* **14**, 5569–5578 (1995).
- Metcalfe, D., Willson, T. A., Hilton, D. J., DiRago, L. & Mifsud, S. Production of hematopoietic regulatory factors in cultures of adult and fetal mouse organs: measurement by specific bioassays. *Leukemia* **9**, 1556–1564 (1995).
- Hilton, D. J., Watowich, S., Katz, L. & Lodish, H. F. Saturation mutagenesis of the WSXWS motif of the erythropoietin receptor. *J. Biol. Chem.* **271**, 4699–4708 (1996).
- Nicola, N. A., Viney, E., Hilton, D. J., Roberts, B. & Willson, T. Cloning of two novel transmembrane ligands for eph-related kinases (LERKs) that are related to LERK2. *Growth Factors* **13**, 141–149 (1996).
- Novak, U. *et al.* Colony-stimulating factor 1-induced STAT1 and STAT3 activation is accompanied by phosphorylation of Tyk2 in macrophages and Tyk2 and JAK1 in fibroblasts. *Blood* **86**, 2948–2956 (1995).
- Naka, T. *et al.* Structure and function of a new STAT-induced STAT inhibitor. *Nature* **387**, 924–929 (1997).

Acknowledgements. We thank R. Simpson and R. Moritz for recombinant mouse IL-6, C. McFarlane for help with production of recombinant cytokines, S. Cory for the pGKpuropA vector, P. Lock for the pGKneo vector and U. Novak for her generous advice regarding electrophoretic mobility shift assays. B. Roberts, D. Cary, L. Di Rago, S. Mifsud and N. Sprigg are thanked for their excellent technical assistance. D.J.H. was supported by a Queen Elizabeth II Postdoctoral Fellowship from the Australian Research Council. T.J.G. is a senior research fellow of the NH&MRC. This work was supported by the Anti-Cancer Council of Victoria, Melbourne, Australia, AMRAD Operations Pty Ltd, Melbourne, Australia, The National Health and Medical Research Council, Canberra, Australia, The J.D. and L. Harris Trust, the National Institutes of Health, Bethesda, Maryland and the Australian Federal Government Cooperative Research Centres Programme.

Correspondence and requests for materials should be addressed to D.J.H. (e-mail: hilton@wehi.edu.au).

A new protein containing an SH2 domain that inhibits JAK kinases

Takaho A. Endo^{*†}, Masaaki Masuhara^{*†}, Masahiro Yokouchi^{*†‡}, Ritsuo Suzuki^{*†}, Hiroshi Sakamoto^{*}, Kaoru Mitsui^{*}, Akira Matsumoto^{*}, Shyu Tanimura^{*}, Motoaki Ohtsubo^{*}, Hiroyuki Misawa^{*}, Tadaaki Miyazaki[§], Nogueira Leonor[§], Tadatsugu Taniguchi[§], Takashi Fujita^{||}, Yuzuru Kanakura[‡], Seturo Komiya[‡] & Akihiko Yoshimura^{*}

^{*} Institute of Life Science, and [‡] Department of Orthopedic Surgery, Kurume University, Aikawamachi 2432-3 Kurume 839, Japan

[§] Department of Immunology, Faculty of Medicine, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

^{||} Department of Tumor Cell Biology, The Tokyo Metropolitan Institute of Medical Science, Bunkyo-ku, Tokyo 113, Japan

[†] Department of Hematology and Oncology, Osaka University Medical School, Suita 565, Japan

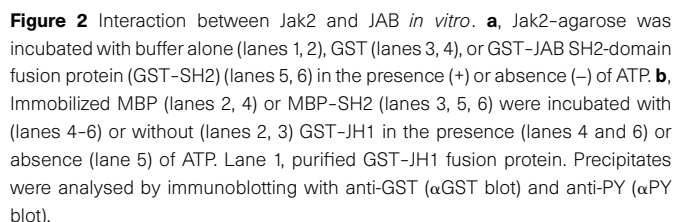
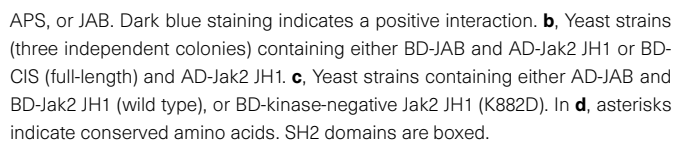
[†] These authors contributed equally to this work.

The proliferation and differentiation of cells of many lineages are regulated by secreted proteins known as cytokines. Cytokines exert their biological effect through binding to cell-surface receptors that are associated with one or more members of the JAK family of cytoplasmic tyrosine kinases. Cytokine-induced receptor dimerization leads to the activation of JAKs, rapid tyrosine-phosphorylation of the cytoplasmic domains, and subsequent recruitment of various signalling proteins, including members of the STAT family of transcription factors, to the receptor complex^{1–5}. Using the yeast two-hybrid system, we have now isolated a new SH2-domain-containing protein, JAB, which is a JAK-binding protein that interacts with the Jak2 tyrosine-kinase JH1 domain⁶. JAB is structurally related to CIS, a cytokine-inducible SH2 protein^{7,8}. Interaction of JAB with Jak1, Jak2 or Jak3 markedly reduces their tyrosine-kinase activity and suppresses the tyrosine-phosphorylation and activation of STATs. JAB and CIS appear to function as negative regulators in the JAK signalling pathway.

We screened several yeast two-hybrid complementary DNA libraries and found a single positive clone, JAB, in a human B-cell library that could interact with the Jak2 JH1 domain (Fig. 1a–c). A database search revealed that the mouse and rat JAB genes are located downstream of the protamine gene cluster⁹. The coding region of the gene for mouse JAB appears to contain no introns, and encodes a 212-amino-acid protein. The JAB protein contains a central SH2 domain (amino acids 79–170) which is most closely related (35% identical) to that of CIS, a cytokine-inducible SH2-containing protein (Fig. 1d) that we have cloned previously⁷.

We examined the specificity of the interaction of JAB with the Jak2 JH1 domain using the yeast two-hybrid assay. We found no evidence for interaction of the Jak2 JH1 domain with other SH2-containing proteins such as Crk, the p85 subunit of phosphatidylinositol-3-OH kinase (PI(3)K), phospholipase C-γ (PLCγ) or a new

We investigated the effect of JAB on JAK-dependent gene activation in three different reporter-gene assay systems: Stat5 activation by erythropoietin (EPO)⁸ (Fig. 4a), Stat3 activation by interleukin (IL)-6 (ref. 12) (Fig. 4b), and *c-fos* promoter activation by IL-2 and IL-3 (ref. 13; Fig. 4c). Transient expression of JAB with the reporter gene inhibited reporter activation in response to both EPO and IL-6 (Fig. 4a,b). IL-2- and IL-3-induced *c-fos* promoter activation was



almost completely inhibited by co-expression of full-length JAB (Fig. 4c). Although the *c-fos* gene may not be activated by STATs, the activation of JAKs is necessary for *c-fos* activation (see the effect of dominant-negative Jak3, J3M1, in Fig. 4c, and ref. 13). Thus, the decrease in *c-fos* activation in this assay therefore probably reflects inhibition of JAKs, rather than STATs, by JAB. Truncated JAB

lacking both the C- and N-terminal domains had little effect on IL-2- and IL-3-induced *c-fos* activation (Fig. 4c, Myc-SH2), indicating that these regions are necessary for full JAK kinase inhibition by JAB. This is consistent with the results shown in Fig. 3e, which indicate that the JAB SH2 domain alone is unable to inhibit JAK activation. Co-expression of JAB did not inhibit activation of the

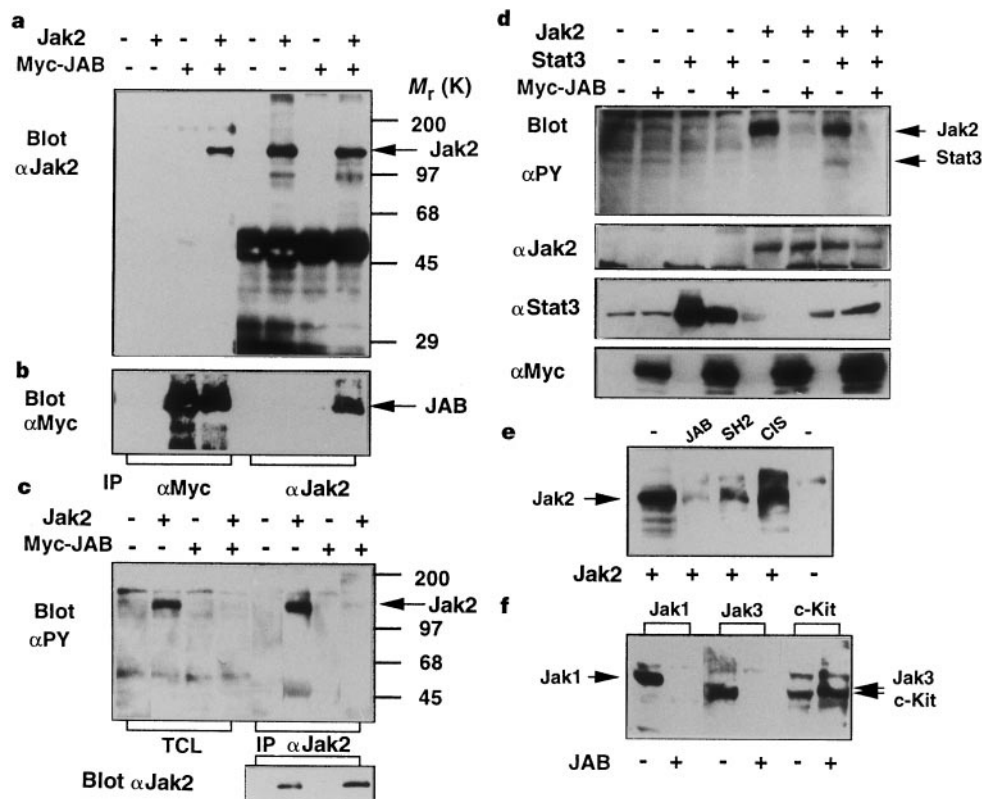


Figure 3 Association of Jak2 with JAB and inhibition of JAK kinase reaction *in vivo*. **a–c**, 293 cells were transfected with Jak2 and/or Myc-JAB constructs, and cell extracts were immunoprecipitated (IP) with antibodies specific for either Myc (IP αMyc) or Jak2 (IP αJak2), then immunoblotted with anti-Jak2 (blot αJak2) (**a**) or anti-Myc (blot αMyc) (**b**) antibodies. **c**, Total cell extracts (TCL) or anti-Jak2 immunoprecipitates (IP αJak2) were blotted with anti-PY (blot αPY) or anti-Jak2 (blot αJak2). **d**, Jak2, Stat3 and JAB constructs were transiently transfected into 293

cells as indicated, and total cell extracts were blotted with anti-PY (αPY), anti-Jak2 (αJak2), anti-Stat3 (αStat3), and anti-Myc (αMyc) antibodies. **e**, 293 cells were transfected with the Jak2 (constructs) (+) or vector (–), either alone or with pcDNA3 carrying full-length JAB (JAB), JAB SH2 domain (SH2), or full-length CIS (CIS). Cell extracts were immunoblotted with anti-PY. **f**, 293 transfected with Jak1, Jak3 or c-Kit(D814V) constructs, either alone (–) or with (+) pcDNA3-Myc JAB. Cell extracts were immunoblotted with anti-PY.

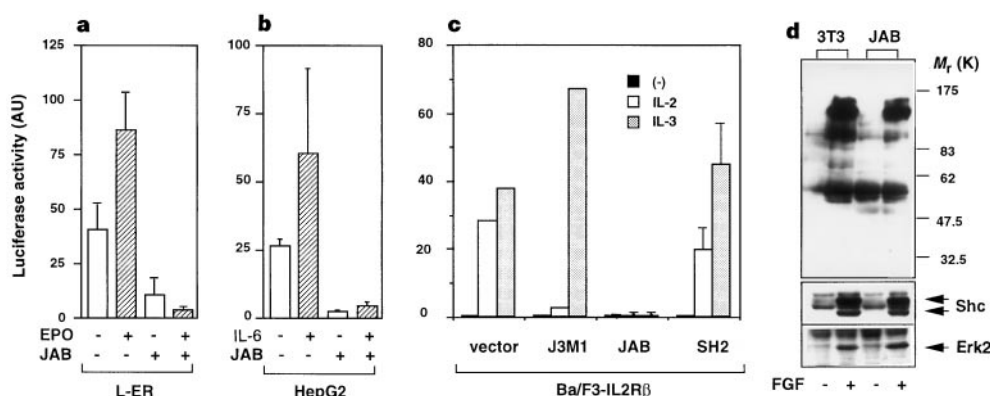


Figure 4 Inhibition of JAK activity by JAB in cells assessed by reporter assay. The indicated cells were transfected with reporter genes for Stat5 (**a**), Stat3 (**b**) or *c-fos* promoter (**c**) with pcDNA3 carrying JAB (+) or JAB, the SH2 domain of JAB (SH2), JH1-truncated Jak3 (J3M1), or vector alone (–). Luciferase activity (AU, arbitrary units) was measured after stimulation without (–) or with (+) 10 unit m^{-1} EPO (**a**),

20 ng ml^{-1} IL-6 (**b**), or either 0.5 nM IL-2 or IL-3 (**c**), for 10–12 h. In **d**, parental NIH3T3 cells and 3T3 cells expressing Myc-JAB (JAB) were stimulated for 10 min either with (+) or without (–) 500 ng m^{-1} FGF. Total cell lysates (upper panel) or anti-Shc or anti-MAP kinase (Erk2) immunoprecipitates (lower panels) were blotted with anti-PY antibody.

c-fos promoter stimulated by cyclic AMP (data not shown), suggesting that JAB does not inhibit any pathways other than those involving JAKs. Furthermore, constitutive overexpression of JAB in NIH3T3 cells did not inhibit basic fibroblast growth factor (FGF)-induced tyrosine-phosphorylation of cellular proteins, including Shc and Erk2 (Fig. 4d), whereas interferon- α -induced antiviral and antiproliferative effects were almost completely blocked in the JAB transformants (data not shown).

We have described the cloning of a cDNA encoding a protein that binds to the catalytic domain of the cytoplasmic tyrosine kinase Jak2. Our results indicate that interaction with JAB inhibits the catalytic activity of JAKs and thus the activation of signalling intermediates such as STATs. CIS also inhibits cytokine signal transduction by competing with Stat5 or other signalling molecules for docking sites on the receptor^{7,8}. The potent inhibition by JAB and CIS of cytokine activity suggests that members of this family may be key regulators of haematopoiesis, of the response to infection and injury, and of development—all biological processes in which JAKs and STATs are important.

Note added in proof: Elsewhere in this issue, JAK inhibitors similar to JAB are reported under the names SOCS¹⁸ and SSI-1 (ref. 19). □

Methods

Two-hybrid screening and cloning of JAB. The two-hybrid screen was done according to ref. 14. The hybrid bait consisted of the Jak2 JH1 domain, and the DNA-binding/dimerization domain of LexA allowed interchain tyrosine phosphorylation of the kinase domain in yeast. The murine Jak2 JH1 domain (amino acids 839–1,127; kindly provided by R. Fukunaga) or the oncogenic c-Kit kinase domain (amino acids 544–975, containing the D814V mutation¹⁰) in a pBTM116 vector was transfected into yeast strain L40. Kinase-negative JH1 was created by substitution of lysine (K) at position 882 to aspartic acid (D) as described¹⁵. A human B-lymphocyte cDNA library in a pACT vector¹⁶ was screened.

Binding of JAB to Jak2 *in vitro*. About 1 μ g purified Jak2 (Upstate Biotechnology) was incubated with 1 μ g purified GST or GST-SH2 fusion protein (amino acids 78–180 of human JAB) in kinase buffer (50 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 5 mM MnCl₂ and 0.1 mM Na₃VO₄) in the presence or absence of 0.2 mM ATP for 2 h at room temperature. After washing twice with phosphate-buffered saline, protein bound to beads was analysed by immunoblotting with anti-phosphotyrosine (PY) (4G10) or goat anti-GST (Pharmacia).

Analysis of JAB/Jak2 interaction in intact cells. Jak2 cDNA in the expression vector pEF-BOS and Myc-epitope-tagged-JAB (Myc-JAB; ref. 17) in pcDNA3 (2.5 μ g per transfection) were transiently expressed in 293 cells grown in 35-mm dishes using calcium phosphate co-precipitation. Cell extracts were immunoprecipitated with either monoclonal anti-Myc antibody (9E10) or polyclonal anti-Jak2 antibody as described⁷. Total cell lysate (50 μ g protein) or the immune complex were further analysed by immunoblotting with anti-PY, anti-Myc, and anti-Jak2. For phosphorylation of Jak1 and Jak3, 293 cells were transfected with 2.5 μ g Jak1 in pCDM8 (from A. F. Wilks), or with Jak3 in pSR α (from T. Shirasawa), or with murine oncogenic c-Kit carrying the D814V mutation in pEF-BOS¹⁰ in the presence or absence of 2.5 μ g JAB cDNA.

Reporter gene assay. Luciferase gene constructs containing an EPO/Stat5-responsive promoter from the CIS gene⁸, an IL-6/Stat3-responsive promoter (based on the acute phase response element¹², or the *c-fos* promoter¹³, were transfected into L929 cells expressing the EPO receptor (L-ER), the hepatocyte cell line HepG2, or into Ba/F3 cells expressing the IL-2 receptor β -chain (Ba/F3-IL2R β), respectively. Luciferase activity was measured as described⁸.

Received 17 March; accepted 2 May 1997.

- Miyajima, A., Kitamura, T., Harada, N., Yokota, T. & Arai, K. Cytokine receptors and signal transduction. *Annu. Rev. Immunol.* **10**, 295–331 (1992).
- Schindler, C. & Darnell, J. E. Jr Transcriptional responses to polypeptide ligands: the JAK–STAT pathway. *Annu. Rev. Biochem.* **64**, 621–651 (1995).
- Ihle, J. N. Cytokine receptor signalling. *Nature* **377**, 591–594 (1995).
- Ihle, J. N. STATs: Signal transducers and activators of transcription. *Cell* **84**, 331–334 (1996).
- Darnell, J. E. Jr, Kerr, I. M. & Stark, G. R. Jak–STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264**, 1415–1421 (1994).

- Harpur, A. G., Andres, A. C., Ziemiecki, A., Aston, R. R. & Wilks, A. F. JAK2, a third member of the JAK family of protein tyrosine kinases. *Oncogene* **7**, 1347–1353 (1992).
- Yoshimura, A. *et al.* A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosine-phosphorylated interleukin-3 and erythropoietin receptors. *EMBO J.* **14**, 2816–2816 (1995).
- Matsumoto, A. *et al.* CIS, a cytokine inducible SH2 protein, is a target of the JAK–STAT5 pathway and modulates STAT5 activation. *Blood* **89**, 3148–3154 (1997).
- Schluter, G. *et al.* Sequence analysis of the conserved protamine gene cluster shows that it contains a fourth expressed gene. *Mol. Reprod. Develop.* **43**, 1–6 (1996).
- Kitayama, H. *et al.* Constitutively activating mutations of c-kit receptor tyrosine kinase confer factor-independent growth and tumorigenicity of factor-dependent hematopoietic cell lines. *Blood* **85**, 790–798 (1995).
- Feng, J. *et al.* Activation of JAK2 catalytic activity requires phosphorylation of Y1007 in the kinase activation loop. *Mol. Cell. Biol.* **17**, 2497–2501 (1997).
- Nakajima, K. *et al.* A central role for Stat3 in IL-6-induced regulation of growth and differentiation in M1 leukemia cells. *EMBO J.* **15**, 3651–3658 (1996).
- Kawahara, A., Minami, Y., Miyazaki, T., Ihle, J. N. & Taniguchi, T. Critical role of the interleukin 2 (IL-2) receptor gamma-chain-associated Jak3 in the IL2-induced *c-fos* and *c-myc*, but not *bcl-2*, gene induction. *Proc. Natl Acad. Sci. USA* **92**, 8721–8728 (1995).
- Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. Mammalian Ras interacts directly with serine/threonine kinase Raf. *Cell* **74**, 205–214 (1993).
- Briscoe, J. *et al.* Kinase-negative mutants of JAK1 can sustain interferon- γ -inducible gene expression but not an antiviral state. *EMBO J.* **15**, 799–809 (1996).
- Harper, J. W., Adams, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinase. *Cell* **75**, 805–816 (1993).
- Turner, D. L. & Weintraub, H. Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev.* **8**, 1434–1447 (1994).
- Starr, R. *et al.* A family of cytokine-inducible inhibitors of signalling. *Nature* **387**, 917–921 (1997).
- Naka, T. *et al.* Structure and function of a new STAT-induced STAT inhibitor. *Nature* **387**, 924–929 (1997).

Acknowledgements. We thank H. Ohgusu for technical assistance, T. Hirano for the APRE/jun promoter construct and STAT3 cDNA, R. Fukunaga for Jak2, A. F. Wilks for the Jak1, T. Shirasawa for rat Jak3 cDNAs, and D. Hilton and R. Starr for reading and editing the manuscript. Part of this work was supported by grants from the Ministry of Science, Education and Culture of Japan, the Kato Memorial Foundation, the Haraguchi Memorial Foundation, the Uehara Memorial Foundation, the Kowa Life Science Foundation, the Motida Memorial Science Foundation and the Naito Memorial Foundation.

Correspondence and requests for materials should be addressed to A.Y. (yosimura@lsi.kurume-u.ac.jp). The nucleotide sequence data reported here will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases under accession numbers AB 000676 (human JAB) and AB 000677 (mouse JAB).

Structure and function of a new STAT-induced STAT inhibitor

Tetsuji Naka*, Masashi Narazaki*, Moritoshi Hirata*, Tomoshige Matsumoto*, Seiji Minamoto*, Atsufumi Aono*, Norihiro Nishimoto†, Tadahiro Kajita‡, Tetsuya Taga§, Kazuyuki Yoshizaki†, Shizuo Akira|| & Tadimitsu Kishimoto*

*Osaka University Medical School Department of Medicine III, 2-2, Yamada-Oka, Suita, Osaka 565, Japan

†Medical Science I, School of Health and Sports Sciences, Department of Osaka University, 2-1, Yamada-Oka, Suita, Osaka 565, Japan

‡Research and Development Center, International Reagents Corporation, 1-2, 1-Chome, Murotani, Nishiku, Kobe 651-22, Japan

§Department of Molecular Cell Biology, Medical Research Institute, Tokyo Medical and Dental University, 2-3-10, Kanda-Surugadai, Chiyoda-ku, Tokyo 101, Japan

||Department of Biochemistry Hyogo College of Medicine, 1-1, Mukogawa, Nishinomiya, Hyogo 663, Japan

The signalling pathway that comprises JAK kinases and STAT proteins (for signal transducer and activator of transcription) is important for relaying signals from various cytokines outside the cell to the inside^{1–3}. The feedback mechanism responsible for switching off the cytokine signal has not been elucidated. We now report the cloning and characterization of an inhibitor of STAT activation which we name SSI-1 (for STAT-induced STAT inhibitor-1). We found that SSI-1 messenger RNA was induced by the cytokines interleukins 4 and 6 (IL-4, IL-6), leukaemia-inhibitory factor (LIF), and granulocyte colony-stimulating factor (G-CSF). Stimulation by IL-6 or LIF of murine myeloid leukaemia cells (M1 cells) induced SSI-1 mRNA expression which was blocked by transfection of a dominant-negative mutant of Stat3,

expressed SSI-1 mRNA in response to IL-4 and G-CSF, respectively. These results show that SSI-1 is induced not only by IL-6 and G-CSF, both of which activate Stat3 (refs 4, 21), but also by IL-4, a cytokine that activates Stat6 (ref. 19). Consistent with these findings, the promoter region of the SSI-1 gene was found to contain Stat3 and Stat6 binding sequences^{4,19,22} (data not shown).

We next examined SSI-1 induction in M1 cells transfected with wild-type Stat3 (M1/Stat3) and a Stat3 mutant (M1/Y705F) (Fig. 2c). Stat3(Y705F) is a dominant-negative form of Stat3, in which a tyrosine at residue 705 that is phosphorylated by a JAK kinase has been substituted by phenylalanine^{6,7}. In this experiment, transfectants containing only the neomycin-resistance gene (M1/Neo) were used as controls. SSI-1 mRNA was more strongly induced by IL-6 plus sIL-6R or LIF in M1/Stat3 cells than in M1/Neo control cells, but was not induced in M1/Y705F cells. These results indicate that the SSI-1 gene is one of the target genes of Stat3 and is induced by the JAK-STAT pathway.

To test the effect of SSI-1 on the gp130-mediated signalling pathway, we established M1 transfectants that constitutively express wild-type SSI-1 (M1/SSI-1) or a mutant SSI-1 (M1/TR) that is

truncated at the C-terminal region which includes the SH2 domain. As shown in Fig. 3a,b, M1/Neo and M1/TR underwent growth arrest and cell death after treatment with IL-6 plus sIL-6R or LIF, as did the parental M1 cells (M1). The dead cells showed features of apoptosis such as chromatin condensation and apoptotic bodies (Fig. 3c). In contrast, growth of M1/SSI-1 cells did not arrest following stimulation (Fig. 3a-c). Expression of the receptor for the immunoglobulin fragment Fc γ (Fc γ R) on days 0, 1, 2 and 3 was examined with flow cytometry in M1/Neo, M1/SSI-1 and M1/TR cells treated with IL-6 plus sIL-6R. Despite a partial inhibition of Stat3 phosphorylation (Fig. 4a), SSI-1 almost completely blocked IL-6-induced Fc γ R expression, which is a direct target gene of Stat3 in M1 cells (Fig. 3d)⁶. These results suggest that SSI-1 inhibits the Stat3-mediated signalling pathway.

Stimulation of cytokines of the IL-6 family induces tyrosine-phosphorylation of the signal-transducing receptor component gp130 to activate Stat3 by JAK kinases²³⁻²⁵. To determine how SSI-1 exerts its inhibitory effect on the gp130-mediated signalling pathway, we measured the tyrosine-phosphorylation of gp130 and Stat3 in M1 cells after stimulation with IL-6 plus sIL-6R. As shown

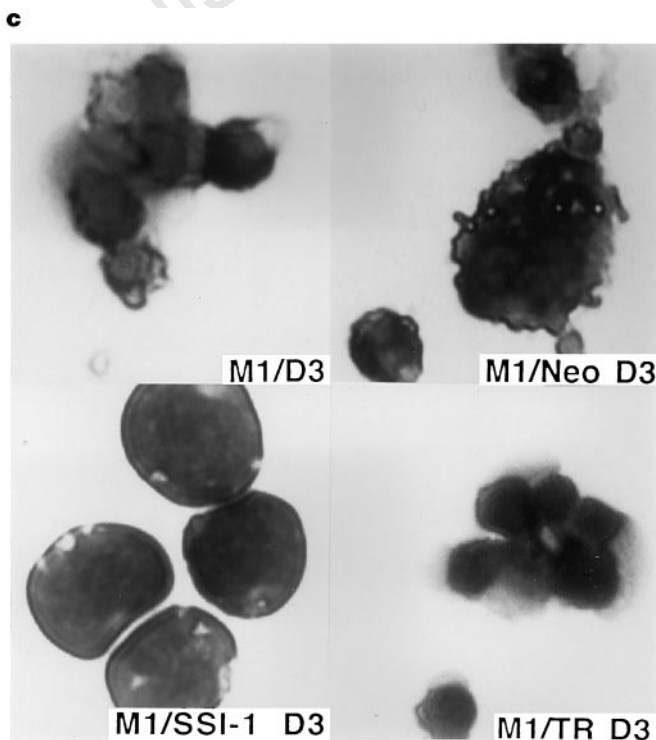
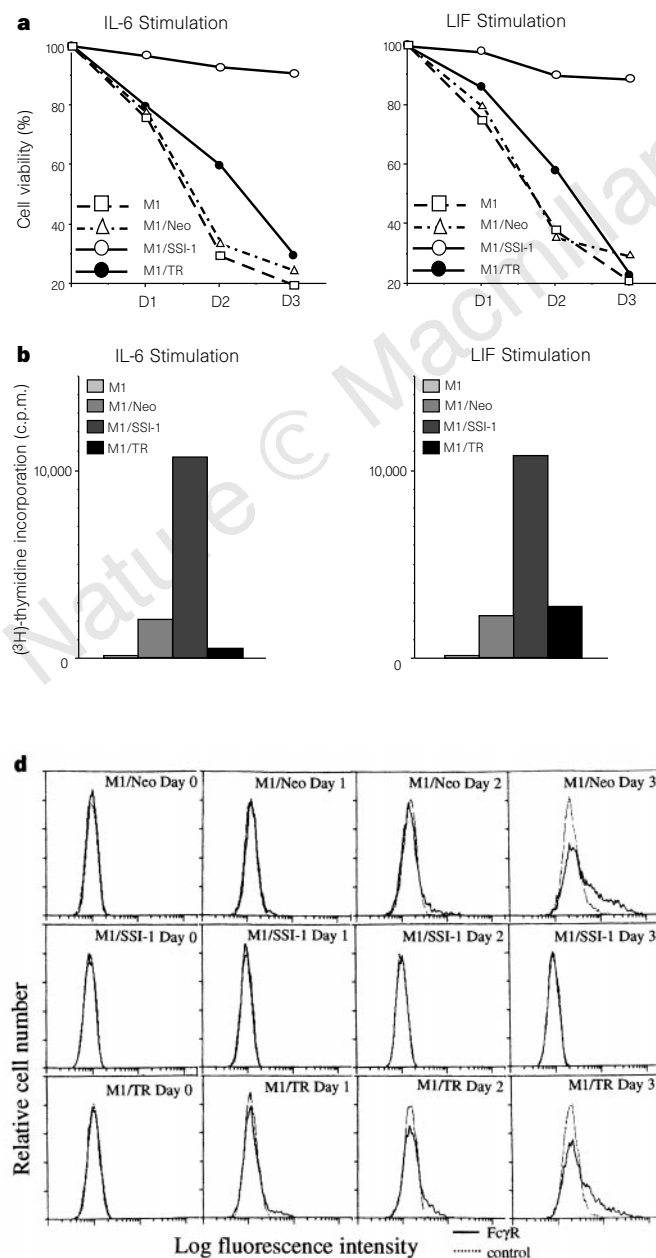


Figure 3 IL-6- or LIF-induced growth arrest and apoptosis of M1 clones. M1, parental M1 cells; M1/Neo, M1 cells carrying a neomycin-resistance gene; M1/SSI-1, M1 cells with wild-type SSI-1; M1/TR, M1 cells carrying mutant SSI-1. **a**, Parental and transfected M1 cells were seeded with IL-6 (300 ng ml⁻¹) plus sIL-6R (500 ng ml⁻¹) or LIF (1,000 U ml⁻¹) at day 0, and the viability of the cells was determined on days (D) 1, 2 and 3. **b**, [³H]thymidine incorporation was measured on day 1 after stimulation. Each column represents the mean of three experiments. **c**, May-Grunwald-Giemsa staining of parental and transfected M1 cells on day 3. **d**, IL-6-induced expression of the Fc γ -receptor, as determined by flow cytometry.

in Fig. 4a, tyrosine-phosphorylation of these molecules was much reduced in M1/SSI-1 cells compared with that in control M1/Neo and M1/TR cells. Next we tested whether SSI-1 interacts directly with JAK kinases. The cDNAs encoding SSI-1 and Jak2 or Tyk2 were co-transfected into COS7 cells; as shown in Fig. 4b, SSI-1 co-precipitated with Jak2 or Tyk2. These results indicate that the activity of JAK kinases is suppressed by SSI-1. To see whether the inhibitory effect of SSI-1 is specific to JAK kinases, we immunoblotted M1/Neo, M1/SSI-1 and M1/TR cell lysates before and after IL-6 plus sIL-6R stimulation with anti-phosphotyrosine. In the case of M1/SSI-1, the intensity of the bands corresponding to gp130 and

Stat3 was significantly reduced compared with the same bands from control M1/Neo and M1/TR cells. The intensity of other bands was comparable among M1/Neo, M1/SSI-1 and M1/TR cell lysates. In addition, there was no difference in the phosphorylation of Flt-3 and the insulin-receptor- β among these cells after stimulation with the corresponding ligands. These results indicate that SSI-1 specifically inhibits the JAK-STAT signalling pathway (Fig. 4c).

Cytokine binding to the receptor induces homo- or heterodimerization of the receptor components and activates JAK kinases that are constitutively associated with the intracellular domains of the receptor components; once activated, these kinases

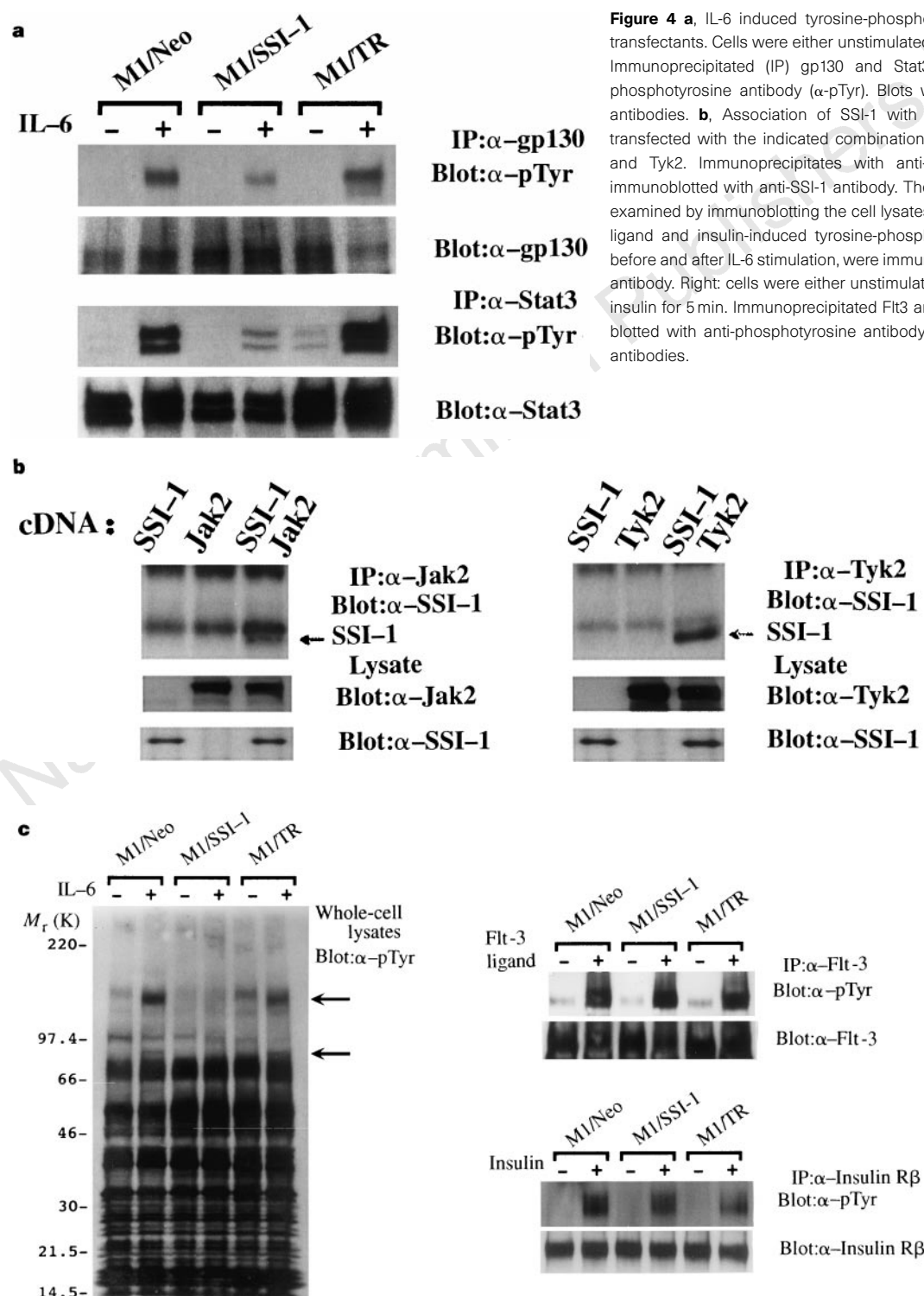


Figure 4 a, IL-6 induced tyrosine-phosphorylation of gp130 and Stat3 in M1 transfectants. Cells were either unstimulated or stimulated with IL-6 plus sIL-6R. Immunoprecipitated (IP) gp130 and Stat3 were immunoblotted with anti-phosphotyrosine antibody (α -pTyr). Blots were reprobed with the respective antibodies. **b**, Association of SSI-1 with Jak2 and Tyk2. COS7 cells were transfected with the indicated combinations of plasmids encoding SSI-1, Jak2 and Tyk2. Immunoprecipitates with anti-Jak2 or anti-Tyk2 antibody were immunoblotted with anti-SSI-1 antibody. The expression of these proteins was examined by immunoblotting the cell lysates. **c**, The effect of SSI-1 on IL-6, Flt-3 ligand and insulin-induced tyrosine-phosphorylation. Left: whole-cell lysates, before and after IL-6 stimulation, were immunoblotted with anti-phosphotyrosine antibody. Right: cells were either unstimulated or stimulated with Flt-3 ligand or insulin for 5 min. Immunoprecipitated Flt3 and insulin-receptor- β were immunoblotted with anti-phosphotyrosine antibody and reprobed with the respective antibodies.

tyrosine-phosphorylate receptor components. STATs then associate with phosphotyrosine residues on the receptor components through their SH2 domains and in turn are tyrosine-phosphorylated by JAK kinases^{1-3,18,25}. Here we have cloned the cDNA encoding SSI-1, an SH2-domain-containing protein that is inducible by Stat3 and inhibits Stat3. It has been reported that CIS participates in negative-feedback regulation of the JAK-STAT signalling pathway^{16,17}. CIS inhibits Stat5 function by directly associating with the Stat5-binding region of the EPO receptor, but the mechanism of SSI-mediated inhibition of the JAK-STAT pathway is quite different. SSI-1 reduces tyrosine-phosphorylation of Stat3 as well as of gp130; as tyrosine-phosphorylation of gp130 is a prerequisite for its association with SSI-1 at the Stat3-binding region, SSI-1 is unlikely to compete with Stat3 for binding to gp130. SSI-1 may block an earlier step in the JAK-STAT signalling pathway than CIS, probably by inhibiting JAK kinases. We have shown that SSI-1 associated with Jak2 and Tyk2, and this molecule has recently been cloned as a JAK-binding protein in a yeast two-hybrid system and also shown to bind Jak1 and Jak3 (A. Yoshimura *et al.*, personal communication). Another point is that SSI-1 is not induced by all STATs, ruling out a role as a general inhibitor of the JAK-STAT pathway.

We have shown that SSI-1 is induced in response to cytokines that act through gp130, and to IL-4. In our experiments, NFS60 and TF-1 cells did not express SSI-1 in response to IL-3 or EPO stimulation, respectively (data not shown). The physiological significance of the cytokine-specific induction of SSI-1 needs to be further investigated. *Note added in proof:* Elsewhere in this issue, JAK inhibitors similar to SSI-1 are reported under the names SOCS²⁶ and JAB²⁷. □

Methods

Preparation of the monoclonal antibody. A hybridoma clone producing the monoclonal antibody (mAb) FL-238, which is reactive against the synthetic oligopeptide (Fmoc) GTFLRLFS (this sequence is highly conserved in the STAT family), was established by fusing spleen cells from BALB/c mice immunized with the synthetic oligopeptide TKPPGTFLRLFSSESSKEG (amino acids 600 to 617 in the SH2 domain of Stat3) coupled to keyhole limpet haemocyanin with mouse myeloma cells P3-X63-Ag8-653. The mAb was purified by protein A affinity chromatography from the ascitic fluid of BALB/c mice.

Isolation of SSI-1 cDNA. Using the monoclonal antibody against the GTFLRLFS motif, SSI-1 cDNA was isolated from a murine thymus cDNA library (lambda ZAP, Stratagene) by using a PicoBlue Immunoscreeing kit.

Cell culture. Myeloid leukaemia M1 cells were cultured in Eagle's minimal essential medium supplemented with double the normal concentrations of amino acids and vitamins and 10% (vol/vol) fetal calf serum (FCS). The IL-6-dependent myeloma MH60 cells were maintained in RPMI 1640 medium supplemented with IL-6 (5 ng ml⁻¹) and 10% FCS. IL-2/IL-4-dependent CT4S cells²⁰ were cultured in RPMI 1640 medium supplemented with IL-4 (10 U ml⁻¹) and 10% FCS. IL-3-dependent myeloid NFS60 cells were maintained in RPMI 1640 medium supplemented with 10% FCS and 10% conditioned medium from the WEHI-3B cell line as a source of IL-3. We had previously constructed dominant-negative forms of Stat3 and established M1 cell lines that constitutively express them⁶. Transfectant M1 cells were cultured with 250 µg ml⁻¹ Geneticin (Gibco).

Northern hybridization. Cells (except M1 cells) were factor-depleted for 4 h in RPMI medium containing 1% BSA, and then all cells were stimulated with cytokines at the following concentrations for various periods: IL-2, 10 ng ml⁻¹; IL-3, 5 ng ml⁻¹; IL-4, 10 U ml⁻¹; IL-6, 100 ng ml⁻¹; sIL-6R, 50 ng ml⁻¹; G-CSF, 20 ng ml⁻¹; LIF, 1,000 U ml⁻¹. Cytoplasmic RNA was extracted using Iso-Gen (Nippon Gene). Total RNA (5 µg ml⁻¹) was electrophoresed on agarose gels and transferred to a nylon membrane (Hybond N⁺, Amersham). The membrane was hybridized with radiolabelled cDNA probes.

Flow cytometry analysis. Transfectants (1 × 10⁵) were cultured with IL-6 (300 ng ml⁻¹) plus sIL-6R (500 ng ml⁻¹) for 3 days. After collection, cells were incubated with 1 µg mouse IgG2a (Capel) for 20 min on ice. After washing with PBS, cells were incubated with fluorescein-isothiocyanate-conjugated anti-mouse IgG antibody (Zymed) for 20 min on ice. Cells were analysed in a fluorescence-activated cell sorter (Becton Dickinson FACS).

Plasmid construction and DNA transfection. Constructs were cloned into pEF-BOS, a mammalian expression vector. Briefly, SSI-1 cDNA was digested with the restriction enzymes *Xba*I and *Pvu*II and then inserted into the *Xba*I site of the pEF-BOS vector (pEF-BOS/SSI-1 (SSI-1)). For the construction of pEF-BOS/SSI-1 (TR), 360 bp of a *Bss*HII-digested fragment was deleted. cDNAs encoding Jak2 and Tyk2 were inserted into the *Xba*I site of the pEF-BOS vector.

M1 cells were transfected by electroporation with the SSI-1(SSI-1/TR) expression vector and pSV2 Neo at a 20:1 ratio, and neomycin-resistant clones were selected in growth medium containing Geneticin (Gibco) at 750 µg ml⁻¹. **Western blot analysis.** Cells were treated with IL-6 (300 ng ml⁻¹) plus sIL-6R (500 ng ml⁻¹), Flt-3 ligand (300 ng ml⁻¹; Genzyme) or insulin (1 µM; Becton Dickinson) for 5 min or were untreated, and were solubilized with lysis buffer (0.5% Nonidet P-40, 10 mM Tris, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM Na₂VO₄) containing protease inhibitors. Immunoprecipitates obtained with anti-gp130 (Upstate Biotechnology), anti-Stat3 (ref. 4), anti-Flt-3 (Santa Cruz Biotechnology) or anti-insulin-receptor-β antibody (Santa Cruz Biotechnology) were resolved by SDS-PAGE under reducing conditions and transferred to nitrocellulose. Filters were probed with anti-phosphotyrosine monoclonal antibody (4G10; Upstate Biotechnology) and re-probed with anti-gp130, anti-Stat3, anti-Flt-3 or anti-insulin-receptor-β antibody, respectively, after stripping the first blots. Blots were visualized with the ECL detection system (Amersham). COS7 cells (1 × 10⁶) were transfected with the designated plasmids encoding SSI-1 (2 µg), Jak2 (20 µg), or Tyk2 (20 µg) by using the DEAE-dextran method. After 48 h, cells were solubilized in lysis buffer and the lysates used for western blot analysis.

Received 26 March; accepted 7 May 1997.

- Darnell, J. E. Jr, Kerr, I. M. & Stark, G. R. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264**, 1415-1421 (1994).
- Ilhe, J. N. Cytokine receptor signalling. *Nature* **377**, 591-594 (1995).
- Kishimoto, T., Akira, S., Narazaki, M. & Taga, T. Interleukin-6 family of cytokines and gp130. *Blood* **86**, 1243-1254 (1995).
- Akira, S. *et al.* Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway. *Cell* **77**, 63-71 (1994).
- Zhong, Z., Wen, Z. & Darnell, J. E. Jr Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science* **264**, 95-98 (1994).
- Minami, M. *et al.* STAT3 activation is a critical step in gp130-mediated terminal differentiation and growth arrest of a myeloid cell line. *Proc. Natl Acad. Sci. USA* **93**, 3963-3966 (1996).
- Nakajima, K. *et al.* A central role for Stat3 in IL-6-induced regulation of growth and differentiation in M1 leukemia cells. *EMBO J.* **15**, 3651-3658 (1996).
- Taga, T. & Kishimoto, T. Cytokine receptors and signal transduction. *FASEB J.* **6**, 3387-3396 (1992).
- Ip, N. Y. *et al.* CNTF and LIF act on neuronal cells via shared signaling pathways that involve the IL-6 signal transducing receptor component gp130. *Cell* **69**, 1121-1132 (1992).
- Kondo, M. *et al.* Functional participation of the IL-2 receptor γ chain in IL-7 receptor complexes. *Science* **262**, 1453-1454 (1994).
- Sakamaki, K., Miyajima, I., Kitamura, T. & Miyajima, A. Critical cytoplasmic domain of the common β subunit of the human GM-CSF, IL-3 and IL-5 receptors for growth signal transduction and tyrosine phosphorylation. *EMBO J.* **11**, 3541-3549 (1992).
- Heldin, C. H. Dimerization of cell surface receptors in signal transduction. *Cell* **80**, 213-223 (1995).
- Adachi, M. *et al.* Mammalian SH2-containing protein tyrosine phosphatases. *Cell* **85**, 15 (1996).
- Yi, T., Mui, A. L.-E., Krystal, G. & Ilhe, J. N. Hematopoietic cell phosphatase associates with the interleukin-3 (IL-3) receptor β chain and down-regulates IL-3-induced tyrosine phosphorylation and mitogenesis. *Mol. Cell Biol.* **13**, 7577-7586 (1993).
- Klingmüller, U., Lorenz, U., Cantley, L. C., Neel, B. G. & Lodish, H. Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. *Cell* **80**, 729-738 (1995).
- Yoshimura, A. *et al.* A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosine-phosphorylated interleukin 3 and erythropoietin receptors. *EMBO J.* **14**, 2816-2826 (1995).
- Mui, A. L., Wakao, H., Kinoshita, T., Kitamura, T. & Miyajima, A. Suppression of interleukin-3-induced gene expression by a C-terminal truncated Stat5: role of Stat5 in proliferation. *EMBO J.* **15**, 2425-2433 (1996).
- Heim, M. H., Kerr, I. M., Stark, G. R. & Darnell, J. E. Jr Contribution of STAT SH2 groups to specific interferon signaling by the Jak-STAT pathway. *Science* **267**, 1347-1349 (1995).
- Hou, J. *et al.* An interleukin-4-induced transcription factor: IL-4 Stat. *Science* **265**, 1701-1706 (1994).
- Hu-Li, J., Ohara, J., Watson, C., Tsang, W. & Paul, W. E. Derivation of a T cell line that is highly responsive to IL-4 and IL-2 (CT4R) and of an IL-2 hyporesponsive mutant of that line (CT4S). *J. Immunol.* **142**, 800-807 (1989).
- Tian, S. S., Lamb, P., Seidel, H. M., Stein, R. B. & Rosen, J. Rapid activation of the STAT3 transcription factor by granulocyte colony-stimulating factor. *Blood* **84**, 1760-1764 (1994).
- Wakao, H., Harada, N., Kitamura, T., Mui, A. L.-E. & Miyajima, A. Interleukin 2 and erythropoietin activate STAT5/MGF via distinct pathways. *EMBO J.* **14**, 2527-2535 (1995).
- Narazaki, M. *et al.* Activation of JAK2 kinase mediated by the interleukin 6 signal transducer gp130. *Proc. Natl Acad. Sci. USA* **91**, 2285-2289 (1994).
- Stahl, N. *et al.* Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6Rβ receptor components. *Science* **263**, 92-95 (1994).
- Stahl, N. *et al.* Choice of STATs and other substrates specified by modular tyrosine-based motifs in cytokine receptors. *Science* **267**, 1349-1353 (1995).
- Starr, R. *et al.* A family of cytokine-inducible inhibitors of signalling. *Nature* **387**, 917-921 (1997).
- Endo, T. A. *et al.* A new protein containing an SH2 domain that inhibits JAK kinases. *Nature* **387**, 921-924 (1987).

Acknowledgements. We thank T. Tanaka for IL-4 and CT4S cells; S. Nagata for pEF-BOS vector and Jak2 cDNA; J. Krolewski for Tyk2 cDNA; K. Yasukawa for recombinant IL-6 and sIL-6R; Y. Shima, H. Danno, K. Kunisada and H. Tagoh for technical assistance; H. Saito for discussion; and A. Nobuhara for secretarial assistance. This work was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture, Japan.

Correspondence and requests for materials should be addressed to T.K. (e-mail: kishimoto@imed3.med.osaka-u.ac.jp). The sequence of SSI-1 has been deposited with Genbank, under accession number AB000710.

Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins

David T. Kirkpatrick & Thomas D. Petes

Department of Biology, Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280, USA

A number of enzymes recognize and repair DNA lesions¹. The DNA-mismatch repair system corrects base–base mismatches and small loops, whereas the nucleotide-excision repair system removes pyrimidine dimers and other helix-distorting lesions. DNA molecules with mismatches or loops can arise as a consequence of heteroduplex formation during meiotic recombination². In the yeast *Saccharomyces cerevisiae*, repair of mismatches results in gene conversion or restoration, and failure to repair the mismatch results in post-meiotic segregation (PMS) (Fig. 1). The ratio of gene-conversion to PMS events reflects the efficiency of DNA repair^{3,4}. By examining the PMS patterns in yeast strains heterozygous for a mutant allele with a 26-base-pair insertion, we find that the repair of 26-base loops involves Msh2 (a DNA-mismatch repair protein) and Rad1 (a protein required for nucleotide-excision repair).

We examined the effects of *msh2* and *rad1* mutations on the meiotic segregation patterns of heterozygous markers in the *HIS4* gene, a recombination hotspot⁵. Msh2 (a homologue of the MutS protein of *Escherichia coli*) is required for the repair of both base–base substitutions and small DNA loops, acting together with either

Msh3 or Msh6 and the MutL homologues Mlh1 and Pms1 (refs 6, 7). Rad1 acts in conjunction with Rad10 to cleave 5' to the pyrimidine dimer during nucleotide-excision repair¹. We used two different types of mutant *his4* alleles: *his4-lopd* (a 26-base-pair (bp) non-palindromic insertion in *HIS4*; ref. 5) and *his4-AAG* (a base-pair change within the initiating ATG codon of *HIS4*; ref. 8). Diploid strains heterozygous for one of these alleles and homozygous for *rad1*, *msh2* or both mutations were sporulated and tetrads were dissected. Spore colonies grown on rich medium were replica-plated to medium lacking histidine and scored as His⁺, His[−] and sectorized His⁺/His[−] (representing PMS events) (Tables 1 and 2).

In DNY27, the wild-type strain heterozygous for *his4-lopd*, only 12% of the aberrant segregation events represent PMS events, indicating that the 26-base loop expected within the heteroduplex is efficiently recognized and repaired. Homozygous *rad1* or *msh2* mutations increase the frequency of PMS about threefold, demonstrating that Rad1 and Msh2 are involved in the repair of large DNA loops. Moreover, they function in the same repair pathway, as the *rad1 msh2* mutant has the same PMS frequency as the *rad1* and *msh2* strains. Surprisingly, the *rad1* mutation also significantly increases the PMS frequency for *his4-lopd* when heterozygous, suggesting that the *RAD1* gene product may be rate-limiting for repair in the *RAD1 MSH2*-dependent pathway. *rad1* and *msh2* mutations have no significant effect on the crossover frequency in the *HIS4* to *LEU2* interval (Table 2).

In a wild-type strain heterozygous for the point mutation *his4-AAG* (PD83), we found that 14% of the aberrant segregants were PMS (Table 2). The *rad1* mutation slightly (but significantly ($P = 0.03$)) increased PMS frequency. In an *msh2* derivative, PMS constituted 87% of the total aberrant segregation events, as expected from previous studies⁹. From these results, we suggest that Rad1 plays a smaller role in the meiotic repair of base–base mismatches than it does in the repair of loops. In previous studies^{10,11}, no effects of the *rad1* mutation on PMS frequency of heterozygous mutations (presumably point mutations) were observed, although crossing-over was reduced in ultraviolet-irradiated *rad1* diploids¹².

The nucleotide-excision repair (NER) system in yeast requires the concerted action of a large number of proteins. The NER endonucleases, Rad1/Rad10 and Rad2, cleave 5' and 3' to the pyrimidine dimer respectively, whereas Rad14 seems to be involved in the recognition of DNA damage¹. To determine whether the effect on the repair of large loops observed with Rad1 required other NER proteins, we examined the effects of *rad2* and *rad14* mutations on

Table 1 Effects of *rad1* and *msh2* mutations on the meiotic segregation patterns of heterozygous markers at the *HIS4* locus

Number of tetrads with various meiotic segregation patterns*																	
Diploid	Mutant <i>his4</i> allele	<i>RAD1</i> †	<i>MSH2</i> †	Other alleles	4:4	6:2	2:6	5:3	3:5	Ab 4:4	7:1	1:7	8:0	0:8	Other PMS	Total tetrads	
DNY27	<i>his4-lopd</i>	+/+	+/+	-	252	54	38	11	1	0	0	1	1	1	0	359	
TP1012	<i>his4-lopd</i>	+/-	+/+	-	201	37	25	10	9	1	0	0	1	0	0	284	
TP1013	<i>his4-lopd</i>	-/-	+/+	-	294	86	28	37	24	3	4	2	3	0	0	481	
DTK223	<i>his4-lopd</i>	+/+	-/-	-	83	21	12	17	4	3	0	0	2	0	0	142	
DTK230	<i>his4-lopd</i>	-/-	-/-	-	61	12	10	13	12	0	2	0	2	0	0	112	
DTK241	<i>his4-lopd</i>	+/+	+/+	<i>rad2</i>	164	22	14	1	4	0	0	0	0	0	0	205	
DTK242	<i>his4-lopd</i>	+/+	+/+	<i>rad14</i>	162	39	18	5	1	0	0	0	2	1	0	228	
MW103‡	<i>his4-Sal</i>	+/+	+/+	-	202	31	50	0	0	0	0	0	2	6	0	291	
DTK257	<i>his4-Sal</i>	-/-	+/+	-	249	66	70	0	0	0	0	0	4	3	0	392	
PD83§	<i>his4-AAG</i>	+/+	+/+	-	142	54	58	11	10	0	1	2	13	23	1	315	
DTK224	<i>his4-AAG</i>	-/-	+/+	-	80	41	48	13	11	4	2	1	10	10	0	220	
No. 5	<i>his4-AAG</i>	+/+	-/-	-	40	6	3	16	25	8	0	1	0	0	12	111	
DTK232	<i>his4-AAG</i>	-/-	-/-	-	40	6	6	22	18	10	1	1	0	1	11	116	

* For all segregation patterns, the first number represents the wild-type allele and the second represents the mutant allele. The segregation patterns include: 4:4 (normal mendelian segregation), 6:2 and 2:6 (gene conversion), 5:3 and 3:5 (tetrads with a single PMS event), Ab 4:4 (aberrant 4:4; one wild-type, one mutant, and two sectorized colonies), 7:1 and 1:7 (tetrads yielding three spore colonies of one genotype and one sectorized colony), 8:0 and 0:8 (tetrads yielding four spores of a single genotype). The 'other PMS' class includes aberrant 6:2 and 2:6 tetrads as well as tetrads with three PMS events.

† Notation: +/+, homozygous wild-type at the designated locus; +/-, heterozygosity for the mutant allele; -/-, homozygosity for the mutant allele.

‡ Ref. 5.

§ Ref. 8.

|| Aberrant segregation patterns from this strain sporulated at 30 °C were examined previously⁹. As here all strains were sporulated at 18 °C, the earlier results are not included.

the segregation of *his4-lop*d (DTK241 and DTK242; Table 1). Neither mutation increased the PMS frequency, indicating that the complete NER complex is not required for the repair of large DNA loops.

As the *his4-AAG* and *his4-lop*d mutations are located about 470 bp apart in *HIS4*, an alternative explanation of our results is that Rad1 affects DNA-mismatch repair as a function of the position of the mutation within the gene. Consequently, we examined the effects of the *rad1* mutation on the aberrant segregation pattern of *his4-Sal*, a 4-bp insertion located at the same position as *his4-lop*d. We found that none of the aberrant segregants was a PMS tetrad (Table 1). As we previously found similar segregation patterns for *his4-Sal* in a *RAD1* strain⁵, we conclude that Rad1 specifically affects the repair of large DNA loops.

There are several other interesting points relating to Tables 1 and 2. First, for strains with the *his4-lop*d mutation, we find that *rad1* increases the frequency of both 5:3 and 3:5 tetrads and reduces the frequency of 2:6 tetrads, but has no effect on the frequency of 6:2 tetrads (compare DNY27 and TP1013). This result suggests that Rad1 is used for restoration repair for loops in which the wild-type strand is donated in forming the heteroduplex (Fig. 1a), because loss

of restoration repair would lead to an increase in 5:3 events without a change in 6:2 events. Rad1 may be used for conversion repair in heteroduplexes in which the mutant strand is donated (Fig. 1b), because loss of conversion repair would lead to both a gain of 3:5 events and a loss of 2:6 events. Thus, Rad1 may be involved in repair in which the wild-type strand is excised from the heteroduplex, because such events are restorations when the wild-type strand is donated and conversions when the mutant strand is donated (Fig. 1). Alternatively, as the sequences at the mismatch depend on which strand is donated, Rad1 may have a sequence-specific role in repair. Note that there is an increase in PMS without a compensating decrease in either the 6:2 or 2:6 classes for the *msh2* and *rad1 msh2* strains containing the *his4-lop*d mutation (Table 2). Either the number of tetrads analysed for these strains was too small to detect an altered ratio of the 6:2 and 2:6 classes, or the roles of Rad1 and Msh2 in repairing loops are slightly different.

Second, in the *msh2* strain heterozygous for *his4-AAG*, the increase in PMS is balanced by an overall decrease in conversion (Table 2). This indicates that Msh2 is required primarily for conversion-type repair for *his4-AAG* and restoration-type repair for *his4-lop*d. This difference probably reflects the nature of the

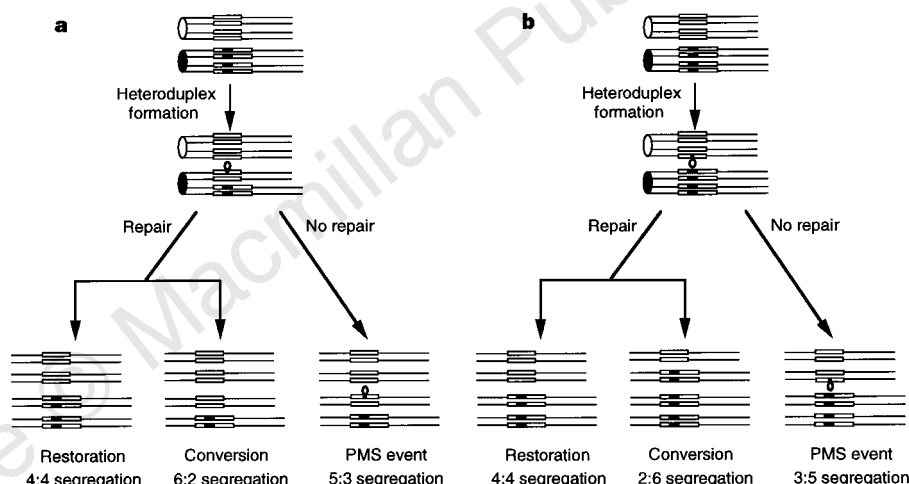


Figure 1 Heteroduplex formation and DNA-mismatch repair during meiosis. Paired meiotic chromosomes are shown, with each chromosome being double-stranded. Centromeres are indicated by black and white ovals; rectangles depict genes with mutant insertions shown in black. **a**, Tetrads in which a wild-type strand is non-reciprocally donated to a mutant gene, or **b**, in which a mutant DNA strand is non-reciprocally donated to a wild-type gene⁴. The resulting hetero-

duplexes contain DNA loops, representing sequences present in the mutant gene but absent in the wild-type gene. Repair of the mismatches can occur either by excision of the loop (followed by resynthesis using the wild-type strand as a template), or by excision of the sequences opposite the loop (followed by resynthesis using the mutant strand as a template). Failure to repair the loop leads to 5:3 or 3:5 postmeiotic segregation (PMS).

Table 2 Effects of *rad1* and *msh2* mutations on the frequencies of aberrant segregation and crossover frequencies in diploid strains heterozygous for *his4-lop*d or *his-AAG*

Relevant diploid genotype (strain)	Ab. seg. (%)	Conversion (%)	PMS (%)	PMS/ab. seg. (%)	<i>HIS4-LEU2</i> distance (cM)
<i>his4-lop</i> d strains					
Wild-type (DNY27)	30 (1.0 ×)	26 (1.0 ×)	4 (1.0 ×)	12 (1.0 ×)	39 (1.0 ×)
<i>rad1/RAD1</i> (TP1012)	29 (0.9 ×)	22 (0.8 ×)	7 (1.8 ×)	24 (2.0 ×)*	33 (0.8 ×)
<i>rad1/rad1</i> (TP1013)	39 (1.3 ×)	25 (0.9 ×)	14 (3.5 ×)**	36 (3.0 ×)**	36 (0.9 ×)
<i>msh2/msh2</i> (DTK223)	42 (1.4 ×)*	25 (0.9 ×)	17 (4.3 ×)**	41 (3.4 ×)**	38 (1.0 ×)
<i>rad1/rad1 msh2/msh2</i> (DTK230)	46 (1.5 ×)**	22 (0.8 ×)	23 (5.8 ×)**	51 (4.3 ×)**	34 (0.9 ×)
<i>his4-AAG</i> strains					
Wild-type (PD83)	55 (1.0 ×)	47 (1.0 ×)	7 (1.0 ×)	14 (1.0 ×)	28 (1.0 ×)
<i>rad1/rad1</i> (DTK224)	64 (1.2 ×)*	50 (1.1 ×)	13 (1.8 ×)*	21 (1.5 ×)	31 (1.1 ×)
<i>msh2/msh2</i> (no. 5)	64 (1.2 ×)	9 (0.2 ×)**	55 (7.9 ×)**	87 (6.2 ×)**	26 (0.9 ×)
<i>rad1/rad1 msh2/msh2</i> (DTK232)	66 (1.2 ×)	12 (0.3 ×)**	53 (7.6 ×)**	82 (5.9 ×)**	46 (1.6 ×)

The data are derived from Table 1. Numbers in parentheses represent the increase relative to wild-type. A single asterisk indicates a statistically significant difference compared to wild type at a *P* value between 0.01 and 0.05; double asterisks indicate a statistically significant difference at a *P* value of less than 0.01. Ab. seg., aberrant segregation.

DNA mismatch (loop as compared to base–base mismatch), because it was previously found that the *msh2* mutation raised PMS and decreased gene conversion for base–base mismatches distributed throughout the *HIS4* gene⁹.

Our results demonstrate a previously undetected requirement for Msh2 for the repair of large loops formed during meiotic recombination. Previous studies showed that *msh2* increased PMS for heterozygous base substitutions and for insertions of 4 bp^{9,13}. Mutations in the *MutL* homologue *MLH1* lead to an increase in PMS for the *MAT* loci (which differ by a large region of heterology) (N. Hunter and R. Borts, personal communication), whereas mutations in *PMS1* (another *MutL* homologue) do not increase PMS frequency for a 38-bp heterology³. Msh2 can bind base–base mismatches and DNA loops *in vitro* in the absence of other mismatch-repair proteins¹⁴, but it is unclear whether other MutS and MutL homologues or an entirely different repair pathway are involved in Msh2-dependent loop repair in meiosis. In mitotic cells, *msh2* mutants seem to be incapable of efficiently correcting loops of 20 bases or larger which result from DNA-polymerase slippage^{15,16}, but these studies would fail to detect small effects.

Other connections between the mismatch-repair and nucleotide-excision-repair systems have been reported. For example, the human MutS α mismatch-repair activity can recognize bulky adducts normally repaired by NER proteins^{17,18}, although repair of these adducts by the mismatch-repair system has not been demonstrated. Conversely, the NER system can repair base–base mismatches, although much less efficiently than pyrimidine dimers¹⁹. With certain direct repeat substrates, *msh2* has been shown to lower mitotic recombination to the same extent as *rad1*, and the *msh2 rad1* mutations have about the same effect as the single mutations, indicating that Msh2 and Rad1 might function in a single pathway²⁰. In addition, Msh2 and Msh3 are required in the Rad1/Rad10-dependent removal of non-homologous tails from double-strand breaks, but have no subsequent role in mitotic DSB-induced recombination (N. Sugawara, F. Paques, M. Colaiacovo and J. Haber, personal communication). Finally, mutations in the *Drosophila* homologue of *RAD1*, *mei-9* (ref. 21), increase PMS²².

Our results indicate that both Rad1 and Msh2 are involved in the meiotic repair of large DNA loops. As we still observe meiotic gene conversion in the *rad1 msh2 his4-lopd* strain, a *RAD1 MSH2*-independent pathway for meiotic repair of DNA loops must exist. For example, some gene conversion events may not involve heteroduplex intermediates. Whatever the mechanism, our results provide evidence for multiple pathways for meiotic gene conversion. □

Methods

Strains and plasmids. All strains are derived from the haploid strains AS13 (*a leu2 ura3 ade6 rme1*) or AS4 (α and *trp1 arg4 tyr7 ade6 ura3 MAL2*)²³. Strain DNY24 is AS13 with the *his4-lopd* (ref. 5) allele, PD73 is AS13 with *his4-AAG*⁸, and MW1 is AS13 with *his4-Sa1* (ref. 5). Strains RKY1452 (AS4 with *msh2*) and RKY1721 (PD73 with *msh2*) were obtained from E. Alani⁹. *rad1* derivatives of haploid strains AS4 (TP1011), PD73 (TP1010), DNY24 (TP1009) and MW1 (DTK256) were constructed by transforming strains with *Bam*HI-treated pDG18 (*rad1::URA3*; provided by R. Schiestl). Ura[−] derivatives of these strains were selected on 5-fluoroorotic acid²⁴, creating strains DTK225, DTK226 and DTK221, respectively. *msh2* derivatives of haploid strains DNY24 (DTK222), DTK225 (DTK229), DTK226 (DTK231) and DTK221 (DTK228) were constructed by one-step integration of *SpeI*–*Bgl*III-digested pII-Tn10LUK *msh2* (ref. 13) (from E. Alani). *rad2* derivatives of AS4 (DTK239) and DNY24 (DTK237) were constructed by one-step transplacement using

SspI-treated pKM55 DNA (from R. Schiestl). *rad14* derivatives of AS4 (DTK240) and DNY24 (DTK238) were constructed by one-step transplacement using *XbaI*/*Bgl*III-treated p14.4 DNA (provided by L. Prakash). The diploids (haploid strains crossed indicated in parentheses) were: DNY27 (ref. 5) (AS4 $\times$ DNY24); PD83 (ref. 8) (AS4 $\times$ PD73); TP1012 (AS4 $\times$ TP1009); TP1013 (TP1011 $\times$ TP1009); DTK223 (RKY1721 $\times$ DTK222); DTK230 (DTK229 $\times$ DTK228); DTK224 (TP1011 $\times$ TP1010); no. 5 (RKY1721 $\times$ RKY1452); DTK232 (DTK229 $\times$ DTK231); DTK257 (TP1011 $\times$ DTK256); DTK241 (DTK239 $\times$ DTK237); DTK242 (DTK240 $\times$ DTK239).

Meiotic analysis. Strains were sporulated at 18°C and dissected using protocols described in ref. 5 (wild-type and *rad* strains) and ref. 9 (*msh2* strains).

Received 12 December 1996; accepted 22 April 1997.

- Friedberg, E. C., Walker, G. C. & Siede, W. *DNA Repair and Mutagenesis* 1–698 (ASM, Washington DC, 1995).
- Modrich, P. & Lahue, R. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **65**, 101–133 (1996).
- Fogel, S., Mortimer, R. K. & Lusnak, K. in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance* (eds Strathern, J. N., Jones, E. W. & Broach, J. R.) 289–339 (Cold Spring Harbor Laboratory Press, New York, 1981).
- Petes, T. D., Malone, R. E. & Symington, L. S. in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis, and Energetics* (eds Broach, J. R., Pringle, J. R. & Jones, E. W.) 407–521 (Cold Spring Harbor Laboratory Press, New York, 1991).
- Nag, D. K., White, M. A. & Petes, T. D. Palindromic sequences in heteroduplex DNA inhibit mismatch repair in yeast. *Nature* **340**, 318–320 (1989).
- Marsischky, G. T., Filosi, N., Kane, M. F. & Kolodner, R. Redundancy of *Saccharomyces cerevisiae* *MSH3* and *MSH6* in *MSH2*-dependent mismatch repair. *Genes Dev.* **10**, 407–420 (1996).
- Johnson, R. E., Kovvali, G. K., Prakash, L. & Prakash, S. Requirement of the yeast *MSH3* and *MSH6* genes for *MSH2*-dependent genomic stability. *J. Biol. Chem.* **271**, 7285–7288 (1996).
- Detloff, P., Sieber, J. & Petes, T. D. Repair of specific base pair mismatches formed during meiotic recombination in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **11**, 737–745 (1991).
- Alani, E., Reenan, R. A. G. & Kolodner, R. D. Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. *Genetics* **137**, 19–39 (1994).
- DiCaprio, L. & Hastings, P. J. Post-meiotic segregation in strains of *Saccharomyces cerevisiae* unable to excise pyrimidine dimers. *Mutat. Res.* **37**, 137–140 (1976).
- Dowling, E. L., Maloney, D. H. & Fogel, S. Meiotic recombination and sporulation in repair-deficient strains of yeast. *Genetics* **109**, 283–302 (1985).
- Resnick, M. A., Game, J. C. & Stasiewicz, S. Genetic effect of UV irradiation on excision-proficient and -deficient yeast during meiosis. *Genetics* **104**, 603–618 (1983).
- Reenan, R. A. G. & Kolodner, R. D. Characterization of insertion mutations in the *Saccharomyces cerevisiae* *MSH1* and *MSH2* genes: evidence for separate mitochondrial and nuclear functions. *Genetics* **132**, 975–985 (1992).
- Alani, E., Chi, N.-W. & Kolodner, R. The *Saccharomyces cerevisiae* Msh2 protein specifically binds to duplex oligonucleotides containing mismatched DNA base pairs and insertions. *Genes Dev.* **9**, 234–247 (1995).
- Sia, E. A., Kokoska, R. J., Dominska, M., Greenwell, P. & Petes, T. D. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. *Mol. Cell. Biol.* **17**, 2851–2858 (1997).
- Tran, H. T., Gordenin, D. A. & Resnick, M. A. The prevention of repeat-associated deletions in *Saccharomyces cerevisiae* by mismatch repair depends on size and origin of deletions. *Genetics* **143**, 1579–1587 (1996).
- Duckett, D. R. et al. Human MutS α recognizes damaged DNA base pairs containing O⁶-methylguanine, O⁶-methylthymine, or the cisplatin-d(GpG) adduct. *Proc. Natl Acad. Sci. USA* **93**, 6443–6447 (1996).
- Mello, J. A., Acharya, S., Fishel, R. & Essigman, J. M. The mismatch repair protein hMSH2 binds selectively to DNA adducts of the anticancer drug cisplatin. *Chem. Biol.* **3**, 579–589 (1996).
- Huang, J.-C., Hsu, D. S., Kazantsev, A. & Sancar, A. Substrate spectrum of human excinuclease: repair of abasic sites, methylated bases, mismatches, and bulky adducts. *Proc. Natl Acad. Sci. USA* **91**, 12213–12217 (1994).
- Saparbav, M., Prakash, L. & Prakash, S. Requirement of mismatch repair genes *MSH2* and *MSH3* in the *RAD1*–*RAD10* pathway of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics* **142**, 727–736 (1996).
- Sekelsky, J. J., McKim, K. S., chin, G. M. & Hawley, R. S. The *Drosophila* meiotic recombination gene *mei-9* encodes a homologue of the yeast excision repair protein Rad1. *Genetics* **141**, 619–627 (1995).
- Carpenter, A. T. C. Mismatch repair, gene conversion, and crossing-over in two recombination-defective mutants of *Drosophila*. *Proc. Natl Acad. Sci. USA* **79**, 5961–5965 (1982).
- Stapleton, A. & Petes, T. D. The Tn3- β -lactamase gene acts as a hotspot for meiotic recombination in yeast. *Genetics* **127**, 39–51 (1991).
- Boeke, J. D., Lacroute, F. & Fink, G. R. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.* **197**, 345–346 (1984).

Acknowledgements. This work was supported by the NIH. We thank R. Borts and J. Haber for communicating unpublished information, and M. Dominska, M. Mears, R. Pukkila-Worley and Q.-Q. Fan for help with tetrad analysis.

Correspondence and requests for materials should be addressed to T.D.P. (e-mail: tompetes@email.unc.edu).

Myc and Ras collaborate in inducing accumulation of active cyclin E/Cdk2 and E2F

Gustavo Leone, James DeGregori, Rosalie Sears, Laszlo Jakoi & Joseph R. Nevins

Nature 387, 422–425 (1997)

Due to a production error in the film output process, essential information was lost from the gel images of Figs 3 and 4 in this Letter. The figures are reproduced here, with legends, and reprints including the corrected figures are available from the authors. □

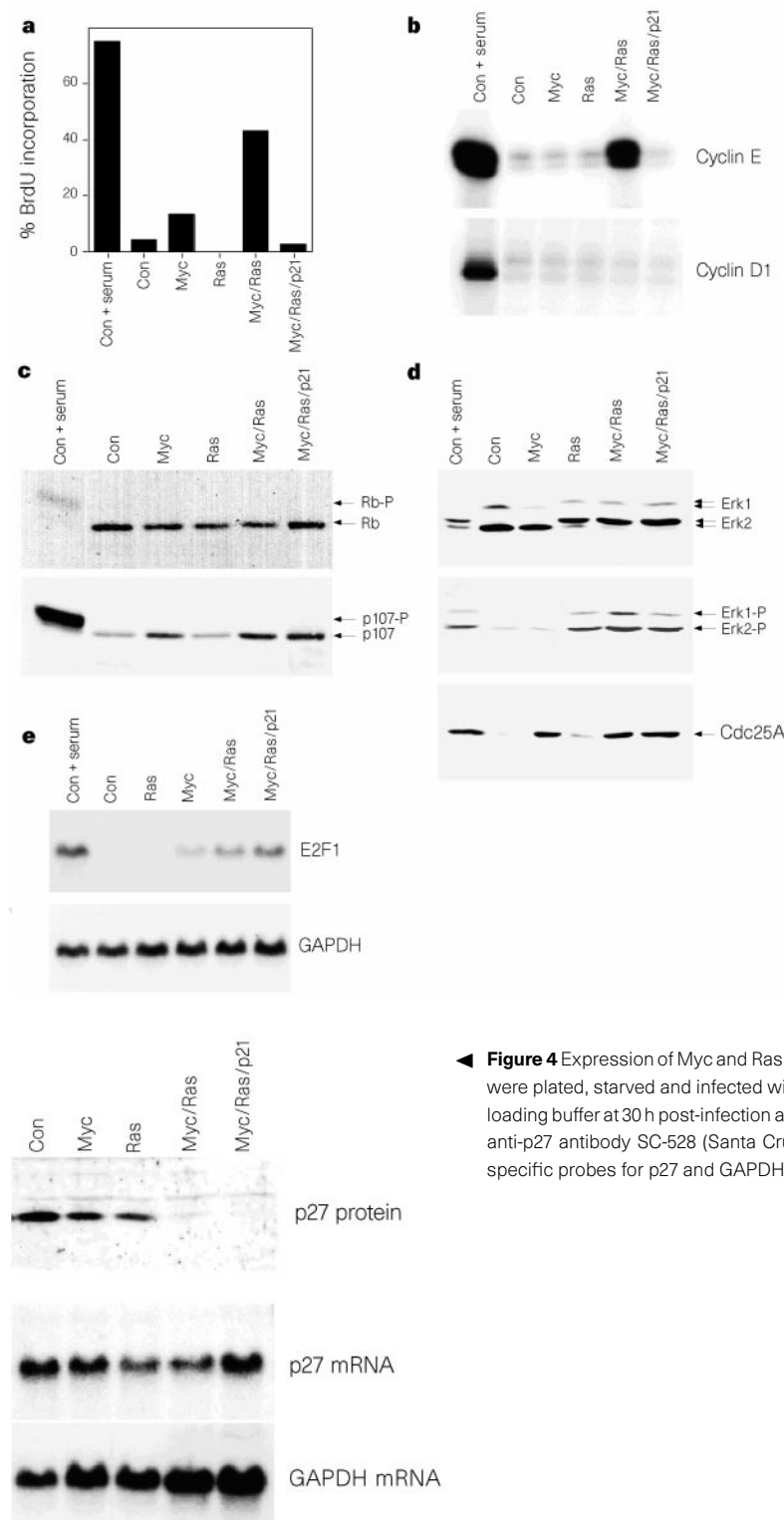


Figure 3 Ras and Myc collaborate to allow accumulation of G1 CDK activity and induction of S phase. **a**, Quiescent REF52 cells were infected with the indicated viruses and BrdU was added after 12 h of infection. Cells were collected 30 h post-infection and BrdU detected by indirect immunofluorescence. Multiplicities of infection were: Ad-Myc (200); Ad-Ras61L (100); Ad-p21 (200); and Ad-Con (300). **b**, Expression of Ras and Myc leads to induction of cyclin E/Cdk2 but not cyclin D/Cdk4 activity. Cells treated as in **a** were collected 30 h post-infection, lysates were immunoprecipitated, and associated kinase activity was determined using histone H1 (top) and GST-Rb (bottom) as previously described²⁷. Affinity-purified rabbit polyclonal antibodies anti-cyclin E (SC-481) and anti-cyclin D1 (SC-450) were purchased from Santa Cruz Biotechnology. **c**, Expression of Ras and Myc does not result in Rb or p107 phosphorylation. Cells treated as in **a** were collected 30 h post-infection in SDS loading buffer and Rb (top) and p107 (bottom) proteins were detected by western blotting using anti-Rb monoclonal antibody 14001A (Pharmingen) and anti-p107 rabbit polyclonal antibody SC-318 (Santa Cruz Biotechnology) as described²⁴. **d**, Coexpression of Ras and Myc does not affect the activity of either protein. Cells were collected 30 h post-infection in SDS loading buffer and Erk1 and Erk2 were detected by western blotting using an anti-Erk1 monoclonal antibody (which cross-reacts with Erk2) M12320 (top) (Transduction Laboratories) or with a phospho-specific rabbit polyclonal anti-Erk1/2 antibody 9101A (middle) (New England Biolabs). The phosphorylated forms of Erk1 and Erk2 represent the upper band of each doublet; the Erk1 protein in the Con + serum sample is present but it is underrepresented in this sample. Cdc25A protein was detected by western blotting using rabbit polyclonal anti-cdc25A antibody SC-97 (bottom) (Santa Cruz Biotechnology). **e**, Myc activates the expression of the *E2F1* gene. Cells treated as in **a** were collected after 15 h infection and processed for northern analysis using specific probes for *E2F1* and GAPDH.

Figure 4 Expression of Myc and Ras leads to a loss of the p27 cyclin-dependent kinase inhibitor. REF52 cells were plated, starved and infected with the indicated viruses as in Fig. 3. Cells were either collected in SDS loading buffer at 30 h post-infection and p27 protein was detected by western blotting using rabbit polyclonal anti-p27 antibody SC-528 (Santa Cruz) (top row) or collected and processed for northern analysis using specific probes for p27 and GAPDH mRNAs (middle and bottom rows).

Finding the meaning in genes

Ways of deciphering and analysing genetic code are the focus in this issue – including clone-design software, a personal server for database searching, a system for site-directed mutagenesis and microsatellite analysis gels.

Synthesized DNA

From Scandinavian Gene Synthesis

GMP-certified, large-scale DNA synthesis

Currently the company manufactures batches of pure DNA of up to 40 bases in length in quantities of up to tens of grams in its GMP-certified facility. However, long oligo-nucleotides of up to 200 bases are also available. The manufacturer provides GMP to the area of DNA so that pharmaceutical and diagnostics companies can have confidence in the high purity and reproducibility of the DNA produced.

Reader Enquiry No. 100

CloneMap

From CGC Scientific

CloneMap is a clone designing and plasmid map drawing program

CloneMap provides an intuitive and interactive environment for fast and easy plasmid map construction, cloning and restriction-site mapping. According to the manufacturer, high-quality plasmid maps can be created with or without a DNA sequence. Map features can be added by restriction-site searching, feature table parsing, or interactive manual entry. Restriction sites can be displayed on the map using predefined enzyme filters, summarized as text files, or listed together with DNA sequences. Free-form graphic drawing objects, floating text and Windows' metafiles can be added to the map to enhance map presentation. Furthermore, map fragments can be inserted or deleted to generate related maps, with concomitant changes in the underlying DNA sequences. CloneMap imports sequences in a variety of formats, such as GenBank, EMBL, GCG, and FastA, and produces publication-quality graphics. CloneMap runs under Windows 3.1, Windows 95 and Windows NT.

Reader Enquiry No. 101

DeCypher II

From Time Logic

A personal workstation/small team server for protein and nucleic acid database searches

This system offers 100 to 300 million affine SW cells per second, a Xilinx reconfigurable computational array and a 200 MHz Pentium-Pro CPU, Windows-NT TCP/IP network sever. Also included is a 1,024 × 768 × 256 high-resolution colour display, a 2–9 gigabyte internal hard drive, a CD-ROM, an optional ×4 CD-ROM recorder and a thin coax and twisted-pair Ethernet port. To improve the economy

Smith–Waterman performance, a dataflow architecture permits supercomputer processing rates while streaming target databases from a conventional low-cost disk subsystem. Employing Xilinx 'software-wired-silicon' reconfigurable, logic-based array processing, the systems adapt to other algorithms through software, eliminating hardware obsolescence associated with older technologies. According to the manufacturer, the personal workstation offers affordable, supercomputer-rate protein and nucleic acid database searches and alignments on the desktop. This product series is suitable for individual, academic or corporate researchers, and for smaller laboratories with a couple of sequencing machines and limited daily query volume. It has been designed to be a secure means of in-house searches of sequences against the public database. It may also prove useful as a fast sequence assembly system by adding up to three accelerators to obtain 300 million affine Smith–Waterman cells per second. Fully loaded, the PW-3 version is comparable to an entry-level Unix add-on, but runs nearly five times faster, according to the manufacturer. Standard Windows-NT networking features ensure integration with your existing Unix, Mac and PC TCP/IP, FTP or NFS heterogeneous network, as well as with Microsoft and Novell networks.

Reader Enquiry No. 102

Electroporating mammalian cells

From EquiBio

Electroporation of in situ adherent cells

This is a system for electroporating mammalian cells without the traditional time-consuming physical and chemical proced-

ures used to detach the cells before electroporation. These *in situ* disposable cuvettes are designed to accept cells grown directly on encapsulated, microporous membranes. The cells remain attached to the membrane throughout the electroporation process, thus reducing cell damage, which in turn improves transfection efficiencies while minimizing cell mortality. This technique effectively bypasses the cell detachment stages, thus reducing the number of practical stages, resulting in an easier and faster transfection procedure.

Reader Enquiry No. 103

In situ hybridization

WCP DNA FISH probes

From Vysis

Offering five assays per kits, this is a second-generation line of WCP whole-chromosome paint DNA probes for all 24 human chromosomes

The fluorescence *in situ* hybridization (FISH) probes have been developed by applying new amplification methods of flow-sorted and micro-dissected chromosomal DNA. This process is said to produce probes that provide brighter fluorescence signals, more uniform chromosome coverage and higher specificity. Higher intensity signals are especially useful for the detection of small translocations that could be missed with the use of probes of lesser intensity. Direct fluorophore labelling provides rapid results in less than one day, according to the manufacturer, without the need for lengthy reagent preparation, detection time or additional detection reagents. In addition, hybridization can be accomplished within four hours. There is also a five-assay kit for low-volume usage. WCP probes have been optimized for research use to detect chromosome translocations, marker chromosomes and human chromosomes in animal/human cell hybrids and to analyse multi-chromosome rearrangement studies, including mutagenesis, radiation and chemical densitometry.

Reader Enquiry No. 104

PNA probes

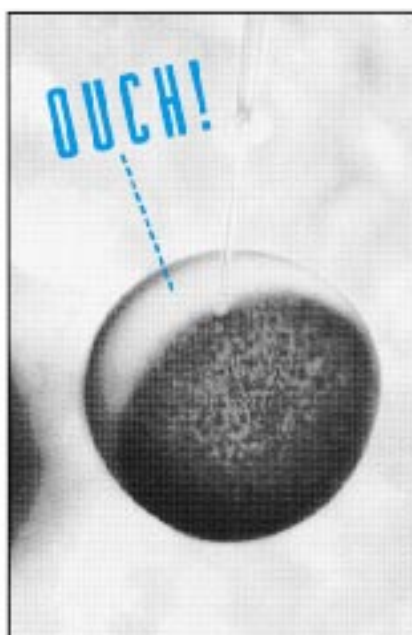
From Dako

PNA (peptide nucleic acid) probes are a new class of probes designed for the detection of specific nucleic acid sequences

These PNA products are all suitable for *in situ* hybridization (ISH), the first of which is the FITC-conjugated Epstein-Barr virus (EBV) PNA probe for the detection of latent



EquiBio's mammalian electroporation system.



Need capped RNA for oocyte injection, transfection, or in vitro splicing studies?

Tired of paying high prices for cap analog and transcription reagents?

The mMESSAGE mMACHINE[®] High Yield Capped RNA Transcription Kit comes complete with 38 (SP6) to 57 (T7 and T3) A₂₀ Units of cap analog – enough for 25 reactions. When you compare the cost for this much cap analog, even at very competitive prices, it's usually more than the cost of the complete mMESSAGE mMACHINE kit. It's like buying cap analog and getting a transcription kit free!

The mMESSAGE mMACHINE kit makes use of Ambion's patented high-yield technology, allowing the routine synthesis of 15–35 µg of capped RNA from 1 µg of template in a 20 µl volume during a 2 hour reaction.

Contact us today for information on our full range of products for RNA Analysis.

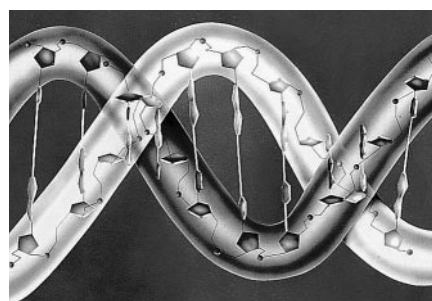
(512) 445-6979 (512) 445-7139 FAX

(800) 888-8804

Ambion[®]

10143

The RNA Company[™]



PNA *in situ* hybridization probes from Dako.

EBV infection and a PNA ISH detection kit for RNA detection. Probes that will be introduced in the future include FITC-conjugated, Epstein-Barr virus (lytic) PNA probe for the detection of active EBV infection and a FITC-conjugated kappa/lambda (mRNA) PNA. All the PNA products are said to be suitable for formalin-fixed, paraffin-embedded tissue sections, frozen sections and cytopins. Enough reagent for at least 40 tests is provided (the number of tests is based on the use of 25 µl of probe and 150 µl of other reagents per tissue section).

Reader Enquiry No. 105

Mutagenesis and mutation detection

MIDAS

From Trevigen

The mutation identification DNA analysis system (MIDAS) is a kit for BRCA1 point mutation detection

Created to assay rapidly the deletion of the 185delAG frameshift point mutation within the BRCA1 gene, the BRCA1 kit is the first in a series of point mutation detection kits for detecting the presence of known point mutations. Also offered is the MIDAScan system for identifying unknown point mutations. The system can be used to detect base-pair mismatches within the target gene and is based upon the annealing of an oligonucleotide probe to a target substrate. If a mismatch is formed during the annealing (reflecting the presence of a point mutation), the probe oligonucleotide is cleaved by the TDG DNA repair enzyme. Cleavage of the oligonucleotide causes the two resulting fragments to 'fall-off' the substrate, leaving the substrate available for another round of probe binding. The gene BRCA1 normally serves as a negative regulator of mammary epithelial cell growth and this function may be compromised in breast cancer either by direct mutation or alterations in gene expression. The BRCA1 gene product is a 190K protein with sequence homology to the granin protein family. Studies of women with early onset familial breast cancer have shown a significant level of mutations within the BRCA1 gene. In particular, a 185delAG frameshift mutation occurs with a high frequency of 0.9 per cent in Ashkenazi Jewish women.

Reader Enquiry No. 106

D Code

From Bio-Rad

A universal mutation detection system that uses a vertical electrophoresis apparatus to perform all gel-based methods

This system uses a modular design that enables customized protocol combinations to be used to increase detection efficiency. It can carry out the five most common methods used for detecting single-base mutations and polymorphisms. In addition, the system offers a new method that matches the detection efficiency of denaturing gradient-gel electrophoresis (DGGE) without the need for a chemical gradient. Temporal, temperature-gradient electrophoresis (TTGE) uses similar principles to DGGE, whereby DNA containing mutations are identified by its melting behaviour. In this case, a denaturing environment, using a constant concentration of urea in the gel, is combined with a temporal temperature gradient. In addition, D_Code offers conventional DGGE capabilities, alongside constant denaturant-gel electrophoresis (CDGE) methods to screen samples rapidly for the presence of previously identified mutations. The system is operated with dedicated MacMelt software, which can predict the melting profile of any DNA sequence up to 3.2 kb, enabling resolution and sensitivity to be enhanced. A cam-operated manual gradient creator produces a linear-gradient gel, which mixes and delivers high- and low-density solutions without using a peristaltic pump or a magnetic stirrer, thus giving reproducible, linear gradients. Other methods that can be run using D_Code include single-strand conformation polymorphism (SSCP), based on the differential migration of denatured wild-type and mutant single-stranded DNA; heteroduplex analysis (HA), based on the differential mobility of re-annealed DNA strands; and protein truncation tests (PTT), where products of transcription/translation reactions are analysed.

Reader Enquiry No. 107

BESS T-Scan

From Epicentre Technologies

A new technology for rapid mutation analysis, detection and localization without sequencing

The manufacturer states that the base-excision, sequence-scanning (BESS) T-Scan kit accurately locates mutations without sequencing single-base changes involving deoxythymidine in up to kilobase lengths of amplified DNA. As a result, the kit can detect and localize greater than 95 per cent of all DNA mutations in a single day, including point mutations, deletions, insertions, repeat expansions and frame shifts. Mutation detection and localization

are obtained simultaneously using a procedure that does not require special gels, gel electrophoresis conditions, hazardous reagents, reannealing steps or extensive reaction optimization. Reagents for PCR and mutation analysis of the amplified DNA are provided in the kit, including MasterAmp PCR enhancer for successful amplification of particularly difficult templates.

Reader Enquiry No. 108

GeneEditor

From Promega

A system formulated for site-directed mutagenesis

This formulation should allow highly efficient, site-directed mutagenesis in any vector carrying the ampicillin resistance gene. Selection is provided by an oligonucleotide that alters the ampicillin resistance gene, increasing its resistance against a number of antibiotics that are provided with the system. In the presence of this antibiotic selection mix, only mutants with increased resistance can grow. By allowing the positive selection of the mutant strand, the desired mutation is achieved at high efficiency, often greater than the stated 90 per cent. The protocol is said to be straightforward, requiring no subcloning or single-stranded DNA template preparation.

Reader Enquiry No. 109

PCR technology

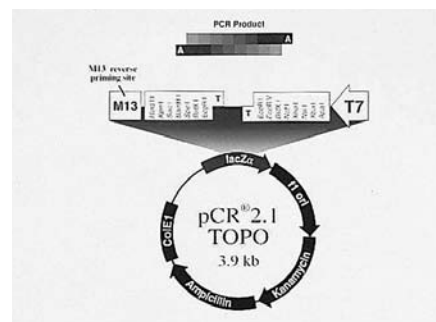
Topo TA

From Invitrogen

A PCR cloning kit that uses a topoisomerase I-activated vector in a five-minute procedure

The kit is said to offer five-minute cloning of *Taq*-generated PCR products with >95 per cent recombinants. The kit uses the pCR 2.1 Topo vector to ligate PCR products in as short as five minutes on the benchtop and without ligase. No special primers, post-PCR modifications or extra steps are needed, states the manufacturer. The kit comes complete with prepared pCR 2.1-Topo vector, PCR reagents, One Shot competent cells, controls and an easy-to-follow manual.

Reader Enquiry No. 110



Topo TA PCR cloning kit uses the pCR 2.1 vector.



Hybaid's oil-free PCR Express thermal cycler.

PCRExpress

From Hybaid

A thermal cycler that has been designed to provide a high level of block performance

A Peltier heat-pump array and temperature-control software are said to give improved speed, uniformity and temperature accuracy. New design features include a range of four interchangeable block options, a new heated lid designed for oil-free operation, interchangeable blocks that utilize automatic block-type recognition and calibration, as well as alphanumeric programming with directory program storage. The system uses an automatically adjusting, heated lid that allows the user to achieve oil-free thermal cycling with tubes and plates from any manufacturer. Active tube-control software on this unit allows users to mimic directly the actual in-sample temperature — stated to be an improved method of in-sample temperature control. This technique allows researchers to monitor the performance of other thermal cyclers, making rapid protocol reoptimizations easy to achieve.

Reader Enquiry No. 111

Electro-Fast

From Advanced Biotechnologies

A new 96-well gel system designed for PCR-based DNA analysis

The Electro-Fast gel casting assembly is said to form sample wells accurately at 9-mm intervals, the standard for 96-well devices. Samples may therefore be loaded from 96-well (or even 384-well) plates using 8-, 12- or 96-channel pipettors. Although not for extended, high-resolution work, this system is said to be well suited for rapid analysis of positive- or negative-outcome PCR (for

example, ARMS-PCR) or for assays with a clear difference in PCR product sizes. Gels are cast directly in the base of the tank, which is transparent to ultraviolet, allowing progress to be monitored during a run without removing the gel.

Reader Enquiry No. 112

Assays and analysis

Atlas Human cDNA Expression Array I

From Clontech Laboratories

A high-density cDNA array representing 588 genes for broad-scale, panoramic expression screening

The array is available now for profiling the expression of hundreds of genes in a single, simple experiment. According to the manufacturer, these genes exhibit tight transcriptional regulation, the hallmark of key players in many biological processes. Oncogenes, apoptosis-related genes, receptors, cell-cycle regulators and other areas of current research interest are represented. Because the array is a nylon membrane, it requires no special equipment to use — it is hybridized and visualized like a Southern blot. Two arrays are shipped with each order so that expression can be compared between two samples, for example, drug-treated and non-treated cells, normal and diseased cells, or between different tissues. This technology should allow researchers to discover in less than a week information that might have previously taken months.

Reader Enquiry No. 113

Express-Check

From American Type Culture Collection

A tool for the inexpensive and quick determination of tissue-specific expression patterns of human genes

The company has announced Express-Check, a kit of DNAs from human tissue-specific cDNA libraries. The kit is a tool for inexpensive and quick determination of tissue-specific expression patterns of human genes.

ADVERTISEMENT



Gramsch Laboratories

New: Endomorphin -1 & -2 Custom Peptide Synthesis:

Shipped within one week:

10 AA DM 400.00; 15 AA DM 520.00;

20 AA DM 660.00; 10-30 mg unpurified;

buy only when satisfied with the HPLC profile

Antiserum Production Service:

you prepare the recombinant protein (in paa-gel or soluble), we produce the antiserum

Distributor for Japan: FUNAKOSHI CO, Tokio 113

Gramsch Laboratories

D-85247 SCHWABHAUSEN, Germany

Phone +49-8138-1769 or +49-8131-5679-0

Fax +49-8138-1260 or +49-8131-5679-24

READER ENQUIRY NO. 95



Express-Check from ATCC uses cDNAs from human expression libraries.

Express-Check can be used to identify a cDNA library containing a sequence of interest before screening the library by hybridization, as well as for the characterization of human gene expression patterns by PCR.

Reader Enquiry No. 114

ECL Plus Western

From Amersham Life Science

A protein detection system designed to increase sensitivity on western blots

This new substrate is said to offer up to a 25-fold increase in sensitivity, providing the ability to detect low-abundance or rare proteins. In addition, the blotting substrate provides a more intense and prolonged pattern of light output; significant levels can be readily detected on PVDF blots up to 48 h after the initial detection stage. Based on the chemiluminescent principle, blotting reagents utilize acridinium ester technology for the detection of horseradish peroxidase-conjugated antibodies. It should also provide a high-sensitivity alternative where ultimate detection levels are required. ECL Plus Western blotting reagents are available in kit format providing sufficient materials for a 1,000-cm² membrane.

Reader Enquiry No. 115

Separation and extraction

Spreadex mini gels

From Guest Elchrom Scientific

Microsatellite analysis submarine gels are composed of a synthetic matrix

According to the manufacturer, Spreadex gels have a resolving power twice that of any self-made or precast polyacrylamide gels. The gels are characterized by an exclusion limit (EL), whereby DNA molecules with lengths exceeding the limit do not migrate in the gels. In the DNA size range from 100 to 300 bp, size differences of 4 bp are resolved on a gel length of 4 cm. Tetranucleotide repeats obtained using Promega's LPL kit for human DNA typing are resolved in a stated 70 min on a Spreadex EL 300 gel. Six different gels are available, with exclusion limits of 300, 400, 500, 600, 800 and 1,200 bp — each is available in five gel formats. Mini-gels have 8- and 12-sample wells, while the wide mini-gels have 26, 50, 52 or 100 sample wells. The wide mini-gels have been designed for high-throughput applications, and 12-channel pipettes are recommended for sample loading. According to the manufacturer, the price per sample of S-100 gels is less than 30 cents. When running the gels in Elchrom's SEA 2000 submarine electrophoresis apparatus, 100–300 samples can be analysed per cycle.

Reader Enquiry No. 116

DNA-Pure genomic kit

From CPG

The kit uses a 1-h protocol for the isolation of high-quality, ultra-pure genomic DNA

With this kit, typical yields for whole blood range from 20 to 50 µg for a 0.3-ml sample; in cell culture yields are 20–50 mg

(2×10^6 cell-sample size); and in 150-mg sizes tissue sample yields are 10–15 µg, states the manufacturer. The economical procedure is said to be easy to follow, efficient and scaleable, eliminating organic extractions completely. Genomic DNA is said to be recovered in high yield and is suitable for restriction enzyme digestion and PCR without further manipulation. The system is based on a special lysis buffer technology and does not use DNA binding resins, thus minimizing DNA shearing. The kit also includes a lysis solution specifically formulated to improve yields with whole blood.

Reader Enquiry No. 117

Nucleon HT kits

From Vector Laboratories

Genomic DNA extraction from hard tissues

The HT kit has been developed by Scotlab Biosciences for high molecular weight DNA extraction from tissues that do not homogenize easily, for example those requiring proteinase K digestion, such as mouse tails, skin and xiphisternum. Previously, obtaining DNA from these samples had involved lengthy protocols, giving unacceptable yields/purity of DNA. The Nucleon matrix that is supplied with these kits removes unwanted proteins without the use of phenol in a simple, rapid protocol. High yields of pure DNA ($A_{260/280}$ ratios ≈ 1.8) are obtained, which are suitable for RFLP and PCR techniques. The kit has sufficient reagents for fifty 25-mg tissue preparations.

Reader Enquiry No. 118

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

BIANORM®
EQUILIBRIUM DIALYSIS SYSTEM

The gentlest method to quantify the binding parameters and thermodynamic data between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

BIANORM
Geräte
P.O. Box 650126
D-81215 München

READER ENQUIRY NO. 92



READER ENQUIRY NO. 92

**CUSTOM ANTIBODIES
TO PEPTIDES AND PROTEINS**

- Purification • Conjugation
- ELISA Titration
- Hybridoma Development

Also Available

1-800-481-9737

e-mail: antibodies@htbio.com
http://www.htbio.com
fax: (760) 788-9694
P.O. Box 1319
Ramona, CA 92065

HTI
HTI Bio-Products, Inc.

READER ENQUIRY NO. 96